KUOPION YLIOPISTON JULKAISUJA E. YHTEISKUNTATIETEET 159
KUOPIO UNIVERSITY PUBLICATIONS E. SOCIAL SCIENCES 159

AUVO RAUHALA

The Validity and Feasibility of Measurement Tools for Human Resources Management in Nursing

Case of the RAFAELA System

Doctoral dissertation

To be presented by permission of the Faculty of Social Sciences of the University of Kuopio for public examination in Auditorium, Mediteknia building, University of Kuopio, on Friday 7th November 2008, at 12 noon

Department of Health Policy and Management
University of Kuopio
Vaasa Central Hospital
Vaasa Hospital District

KUOPION YLIOPISTO

KUOPIO 2008

Distributor: Kuopio University Library
P.O. Box 1627
FI-70211 KUOPIO
FINLAND
Tel. +358 40 355 3430
Fax +358 17 163 410
http://www.uku.fi/kirjasto/julkaisutoiminta/julkmyyn.html

Series Editors: Jari Kylmä, Ph.D.
Department of Nursing Science

Markku Oksanen, D.Soc.Sc. Philosophy
Department of Social Policy and Social Psychology

Author's address: Vaasa Hospital District
Conservative Division
FI-65130 VAASA
FINLAND

Supervisors: Professor Juha Kinnunen, Ph.D.
Department of Health Policy and Management
University of Kuopio

Professor Olli-Pekka Ryynänen, D.Med.Sc.
University of Kuopio
Department of Public Health and Clinical Nutrition
Unit of Family Medicine
Kuopio University Hospital
Division of General Practice

Reviewers: Professor Arja Isola, Ph.D.
Department of Health Sciences
University of Oulu

Professor Kari Lindström, Ph.D.
Finnish Institute of Occupational Health
Helsinki

Opponent: Professor Marja Blom, Ph.D.
Division of Social Pharmacy
Faculty of Pharmacy
University of Helsinki

ISBN 978-951-27-0818-5
ISBN 978-951-27-1069-0 (PDF)
ISSN 1235-0494

Kopijyvä
Kuopio 2008
Finland

Rauhala, Auvo. The validity and feasibility of measurement tools for human resources management in nursing – Case of the RAFAELA system. Kuopio University Publications E. Social Sciences 159. 2008. 126 p.
ISBN 978-951-27-0818-5
ISBN 978-951-27-1069-0 (PDF)
ISSN 1235-0494

ABSTRACT

Background of the study: Nurses' increasing workload, their job dissatisfaction and the lack of nurses have emphasised the significance of the central aims of human resource management (HRM): the simultaneous occurrence of high performance and high occupational welfare. Numerous patient classification systems (PCSs) have been developed for the estimation of nurses' workload, but only relatively few of them have been scientifically validated. The RAFAELA PCS has become widely used in Finland. It uses the OPC (Oulu Patient Classification) for measuring nursing intensity and the PAONCIL (Professional Assessment of Optimal Nursing Care Intensity Level) to allow an analysis of nurses' workload in relation to an optimum. However, the validity and feasibility of this PCS has not been previously sufficiently determined.

Aim of the study: This study focused on evaluating whether the RAFAELA system was valid and feasible enough to be used as a measurement tool for HRM in nursing in the wards of somatic specialized health care in Finland. HRM was used as one theoretical approach of this thesis.

Data and methods: The materials consisted of Finnish secondary health care hospitals' results with the OPC and PAONCIL in routine use. In the first two sub-studies, the material comprised the daily OPC and PAONCIL results of eight wards at one hospital in 1996–1997. In the third sub-study, data on the optimal nursing intensity analyses were gathered from 61 wards from eight hospitals in 1997–2001. In the fourth sub-study, the PAONCIL values and 'non-patient' questions, concerning factors unrelated to the nursing activities of patients, from 4 870 questionnaires and the corresponding OPC values were gathered in 22 wards of a hospital in 2002. In the fifth sub-study, nurses' workload as a ratio of observed to optimal nursing intensity per nurse and sick leave data were gathered from five hospitals: 31 wards and 877 nurse employees in 2004. The statistical methods were mainly linear regression analyses. The negative binomial regression model was used in the fifth sub-study.

Results: In the first and second sub-studies, the determination of optimal nursing intensity could be successfully performed in five out of eight wards for adults (concurrent validity). The PAONCIL was used as a 'gold standard', and the explanatory ratio was 37%. The optimal OPC per nurse values were close to each other in five out of six wards for adults, ranging between 20.2 – 24.3 points per nurse. The distribution of the PAONCIL scores was similar to the normal distribution, but the 0-value was to some extent over-represented. The PAONCIL value was independent of the amount of nurses' working years and of different occupational categories. In the analysis of the construct validity of the OPC, the explanatory power of the linear regression analysis did not increase when several independent variables were added. In the third sub-study, the results of the analysis of optimal nursing intensity could be regarded as reliable if the PAONCIL response rate was above 70%, the period of the PAONCIL examination at least 3 – 4 weeks, the mean PAONCIL value < 0.65 and the explanatory power > 25 per cent. In the fourth sub-study, 'non-patient' questions were answered in 26% of 4 870 questionnaires. Eight questions were grouped into four factors: administration; staff resources and mental stress; co-operation within and between units. The explanatory power between OPC and PAONCIL had a median of 45 per cent. Including the non-patient questions raised this to 55 per cent. In the fifth study, the mean workload was 9% (SD=18%) above the optimum. There was a linear trend between increasing workload and increasing periods of sick leave ($p \leq 0.006$). Among nurses with a workload ≥30% above the optimum, the rate of self-certified periods of sick leave was 1.44 (95% CI 1.13-1.83) times higher than among those with an optimum workload. The corresponding rate ratio for medically certified sick leave was 1.49 (1.10-2.03). These excess rates of sickness absence resulted in 12 extra sick leave days per person-year. About 5-6% of the increased productivity resulting from work overload was lost in increased periods of sick leave.

Conclusions and recommendations:

The OPC measure can be considered a well-validated measure of nursing intensity. It measures patient-associated workload in specialized health care wards sufficiently reliably for HRM purposes. The PAONCIL measure and the whole RAFAELA system were proved to be valid. The determination of optimal nursing intensity can be performed successfully in most wards. The limits of the RAFAELA system have now been thoroughly determined. The OPC and the PAONCIL were shown to measure the same phenomenon. The increased workload of nurses was associated with their markedly increased sick leave. Work overload may significantly increase occupational health problems among employees. The RAFAELA system was shown to have many features that can support HRM in nursing.

National Library of Medicine Classification: WY 105
Medical Subjects Headings: Nursing; Workload; Personnel Staffing and Scheduling; Classification; Validation Studies; Sick Leave; Regression Analysis

Rauhala, Auvo. Validitet och genomförbarhet av mätningsinstrument för ledning av mänskliga resurser i vård. Case RAFAELA system. Kuopio Universitets Publikationer E. Samhällsvetenskap 159. 2008. 126 s.
ISBN 978-951-27-0818-5
ISBN 978-951-27-1069-0 (PDF)
ISSN 1235-0494

SAMMANDRAG

Undersökningens bakgrund: Skötarnas ökade arbetsmängd och missnöje med arbetet, samt rådande brist på skötare, har betonat vikten av de centrala målen i ledningen av mänskliga resurser (HRM): hög prestation och högt välmående i arbetet samtidigt. Talrika system för patientklassificering har utvecklats för att uppskatta skötarnas arbetsmängd, men endast ett fåtal system har genomgått en vetenskaplig validering. RAFAELA-systemet för patientklassificering används i stor utsträckning i Finland. Det använder OPC (Oulu Patient Classification) för mätning av vårdtyngd och PAONCIL (Professional Assessment of Optimal Nursing Care Intensity Level) för analys av skötarnas arbetsmängd i förhållandet till optimum. Systemets validitet och tillämpning har emellertid inte tidigare tillräckligt definierats.

Undersökningens syfte: Denna undersökning fokuserade på evalueringen av RAFAELA-systemets tillräckliga validitet och användbarhet vid användning som en mätinstrument för skötarnas HRM på bäddavdelningar inom specialsjukvården i Finland. HRM användes som teoretisk utgångspunkt i denna avhandling.

Material och metodik: Allt material som använts har bestått av rutinmässiga bestämningar med OPC och PAONCIL inom den somatiska specialsjukvården i Finland. I de två första delstudierna bestod materialet av 8 avdelningar på ett sjukhus åren 1996-1997. Dagliga OPC och PAONCIL-resultat insamlades. I den tredje delstudien insamlades data från analyser av optimal vårdtyngd från 61 avdelningar på 8 sjukhus åren 1997 – 2001. I den fjärde delstudien insamlades PAONCIL-värden och 'icke-patient'- (gällande faktorer som inte tangerade patienternas vårdaktiviteter) svar från 4 870 frågeformulär, samt motsvarande OPC-värden från 22 avdelningar på ett sjukhus år 2002. I den femte delstudien insamlades data angående skötarnas arbetsmängd, uttryckt som förhållandet mellan observerad och optimal vårdtyngd samt data om sjukledigheter från fem sjukhus. Delstudien innefattade 31 avdelningar och 877 skötare år 2004. Vid den statistiska analyseringen av materialet användes främst lineära regressionsanalyser. I den femte delstudien användes en negativ binomial regressionsmodell.

Resultat: I de två första delstudierna kunde den optimala vårdtyngden fastställas på fem av åtta vuxenavdelningar (samtidig validitet). PAONCIL användes som 'gyllene standard'. Variationen i vårdtyngd, mätt som mängden OPC-poäng per skötare, förklarade 37 % av variationen i PAONCIL-värdena. Den optimala mängden OPC-poäng per skötare var nära varandra, mellan 20,2 och 24,3 poäng per skötare, på fem av sex vuxenavdelningar. PAONCIL-poängen var normalfördelade, förutom den något överrepresenterade andelen 0-värden. PAONCIL-värdet var oberoende av skötarnas arbetsår och typ av skötare. I begreppsvaliditetsanalys av OPC ökade inte den lineära regressionsanalysen förklarningsgraden, fastän man tillsatte flera oberoende variabler. I den tredje delstudien kunde resultaten av den optimala vårdtyngden anses som tillförlitliga, såvida PAONCIL-svarsprocenten var > 70 %, PAONCIL-undersökningsperioden ≥ 3 – 4 veckor, PAONCIL-värdet i medeltal < 0,65 och förklarningsgraden > 25 %. I den fjärde delstudien hade man svarat på 'icke-patient'-frågorna i 26 % av 4870 frågeformulär. Åtta frågor grupperades i fyra faktorer: administration; personalresurser och mental stress; samarbete inom och mellan enheter. Medianen på förklarningsgraden mellan OPC och PAONCIL var 45 %. När 'icke-patient'-faktorer inkluderades i analyserna, ökade andelen till 55 %. I den femte delstudien var den genomsnittliga arbetsmängden i medeltal 9 % (SD=18%) över optimum. Det fanns ett lineärt samband mellan ökad arbetsmängd och ökat antal sjukledigheter ($p \leq 0.006$). Hos skötare med arbetsmängd ≥30% över optimum var antalet självanmälda sjukledigheter 1,44 (95 % CI 1,13-1,83) gånger högre än hos skötare med optimal arbetsmängd. Det motsvarande antalet sjukledigheter, dokumenterade med läkarintyg, var 1,49 (1,10-2,03) gånger högre. Detta ökade antal sjukledigheter ledde till 12 extra sjukledighetsdagar per personår. Cirka 5-6 % av den ökade produktiviteten, som ett resultat av ökad arbetsmängd, förlorades i form av ökade sjukledigheter.

Konklusioner och rekommendationer: OPC-mätaren kan anses som en grundligt validerad vårdtyngdsmätare. Den mäter den till patientvården associerade arbetsmängden, med tillräcklig tillförlitlighet för HRM ändamål på bäddavdelningar inom specialsjukvården. Även PAONCIL mätaren och hela RAFAELA-systemet har visat sig vara valida. Det är möjligt att fastställa den optimala vårdtyngden på de flesta avdelningar. RAFAELA-systemets gränser har nu blivit grundligt fastställda. Man kunde påvisa, att OPC och PAONCIL mäter samma fenomen. Den ökade arbetsmängden hos skötarna hade samband med en märkbar ökning av deras sjukledigheter. En för stor arbetsmängd kan märkbart öka personalens arbetsrelaterade hälsoproblem. RAFAELA-systemet visade sig innehålla många särdrag till stöd för HRM bland skötare.

Rauhala, Auvo. Hoitajien inhimillisten voimavarojen johtamisen mittausvälineiden validius ja käyttökelpoisuus. Case RAFAELA-järjestelmä. Kuopion yliopiston julkaisuja E. Yhteiskuntatiede 159. 2008. 126 s.
ISBN 978-951-27-0818-5
ISBN 978-951-27-1069-0 (PDF)
ISSN 1235-0494

TIIVISTELMÄ

Tutkimuksen tausta: Hoitajien lisääntyvä työmäärä, heidän tyytymättömyytensä työhön ja hoitajapula ovat korostaneet ihmisvoimavarojen johtamisen (HRM) keskeisten päämäärien tärkeyttä: erinomainen työsuoritus ja työtyytyväisyys samanaikaisesti. Hoitajien työmäärän arvioimiseksi on kehitetty lukuisia hoitoisuusluokitusjärjestelmiä, mutta vain suhteellisen harvat niistä on validoitu tieteellisesti. RAFAELA-järjestelmän käyttö on muodostunut laajaksi Suomessa. Se käyttää OPC:tä (Oulu Patient Classification) hoitoisuuden mittaamiseen ja PAONCIL:ia (Professional Assessment of Optimal Nursing Care Intensity Level) hoitajien työmäärän vertaamiseksi optimiin. Järjestelmän validiteettia ja soveltuvuutta ei kuitenkaan ole aiemmin riittävästi määritetty.

Tutkimuksen päämäärät: Tämän tutkimuksen päämääränä oli evaluoida, oliko RAFAELA-järjestelmä riittävän validi ja käyttökelpoinen, jotta sitä voidaan Suomen erikoissairaanhoidon osastoilla käyttää hoitajien henkilöstövoimavarojen mittausvälineenä. HRM:a käytettiin yhtenä teoreettisena lähestymistapana tutkimukseen.

Aineisto ja menetelmät: Käytetyt aineistot koostuivat suomalaisten somaattisen erikoissairaanhoidon laitosten rutiinikäytössä olevilla OPC- ja PAONCIL-mittareilla saaduista tuloksista. Kahdessa ensimmäisessä osatyössä aineistot käsittivät yhden sairaalan 8 osastoa vuosina 1996–1997. Kerättiin päivittäiset OPC- ja PAONCIL-tulokset. Kolmannessa osatyössä kerättiin optimaalisen hoitoisuuden määrittämisen tiedot 61 osastolta kahdeksasta sairaalasta vuosina 1997 – 2001. Neljännessä osatyössä kerättiin 4870 kyselykaavakkeesta PAONCIL-arvot ja vastaukset kysymyksiin, jotka koskivat muusta kuin potilashoidosta aiheutunutta työmäärää sekä vastaavat OPC-arvot yhden sairaalan 22 vuodeosastolta v. 2002. Viidennessä osatyössä kerättiin 5 sairaalasta aineisto hoitajien työmäärästä, ilmaistuna hoitajaa kohti lasketun havaitun ja optimaalisen hoitoisuuden suhteena. Mukana oli 31 osastoa ja 877 hoitajaa v. 2004. Analyysit olivat pääosin lineaarisia regressioanalyysejä. Viidennessä osatyössä käytettiin negatiivisen binomijakautuman regressiomallia.

Tulokset: Ensimmäisessä ja toisessa osatyössä optimaalinen hoitoisuus kyettiin määrittämään onnistuneesti viidellä kahdeksasta aikuisosastosta (yhtäaikaisvaliditeetti). PAONCIL:ia käytettiin 'kultaisena standardina' ja selitysosuus oli 37%. Optimaaliset OPC-pistemäärät hoitajaa kohti olivat lähellä toisiaan 5:llä aikuisosastolla 6:sta, välillä 20,2 – 24,3 pistettä per hoitaja. PAONCIL-pisteet jakautuivat normaalijakauman tavoin, paitsi että 0-arvo oli jossakin määrin yliedustettuna. PAONCIL-arvo ei riippunut hoitajien työvuosista eikä ammattiryhmästä. OPC:n rakennevaliditeetin analyysissä lineaarisen regressioanalyysin selitysosuus ei lisääntynyt siitä, että useita lisämuuttujia sijoitettiin riippumattomiksi muuttujiksi. Kolmannessa osatyössä optimaalisen hoitoisuuden analyysien tuloksia voitiin pitää luotettavina, mikäli PAONCIL-vastausprosentti oli yli 70 %, PAONCIL-tutkimusjakso kesti vähintään 3 – 4 viikkoa, keskimääräinen PAONCIL-arvo oli alle 0,65 ja selitysosuus oli yli 25 %. Neljännessä osatyössä lisäkysymyksiin, jotka koskivat muusta kuin potilashoidosta aiheutunutta työmäärää oli vastattu 26 %:ssa 4870 kaavakkeesta. Kahdeksan kysymystä ryhmittyivät neljäksi faktoriksi: hallinto; henkilöstöresurssit ja henkinen stressi; yksikköjen sisäinen ja niiden välinen yhteistyö. OPC:n ja PAONCIL:in välisen selitysosuuden mediaani oli 45 %. Edellä mainittujen lisäkysymysten sisällyttäminen analyysiin nosti selitysosuuden 55 %:iin. Viidennessä osatyössä keskimääräinen työmäärä ylitti optimin 9 %:lla (SD=18 %). Lisääntyvän työmäärän ja lisääntyvien sairauslomien määrän välillä oli lineaarinen yhteys ($p<0,006$). Niillä hoitajilla, joiden työmäärä ylitti optimin ≥30 %, itse ilmoitettujen sairauslomien määrä oli 1,44 (95 % CI 1,13–1,83) kertaa korkeampi kuin hoitajilla, joiden työmäärä oli optimaalinen. Vastaava suhde lääkärintodistuksella osoitettujen sairauslomien osalta oli 1,49 (1,10–2,03). Nämä lisääntyneet sairauslomat johtivat 12 ylimääräiseen sairauslomapäivään henkilövuotta kohti. Noin 5-6 % suuren työmäärän aiheuttamasta kohonneesta tuottavuudesta menetettiin lisääntyneinä sairauslomina.

Johtopäätökset ja suositukset: OPC:tä voidaan pitää monipuolisesti validoituna hoitoisuusmittarina. Se mittaa riittävän luotettavasti HRM-luonteiseen käyttöön potilashoitoon liittyvää työmäärää erikoissairaanhoidon vuodeosastoilla. Myös PAONCIL ja koko RAFAELA-järjestelmä osoitettiin valideiksi. Optimaalisen hoitoisuuden määritys onnistuu useimmilla osastoilla. RAFAELA-järjestelmän rajat ja reunaehdot on nyt perusteellisesti selvitetty. OPC- ja PAONCIL-mittarien osoitettiin mittaavan kutakuinkin samaa ilmiötä. Hoitajien lisääntynyt työmäärä oli yhteydessä heidän sairauslomiensa merkittävään lisääntymiseen. Liiallinen työmäärä saattaa lisätä henkilöstön työterveydellisiä ongelmia merkittävästi. RAFAELA-järjestelmän osoitettiin sisältävän monia piirteitä, jotka tukevat HRM-perusteista hoitajien johtamista.

Yleinen suomalainen asiasanasto: hoitotyö; hoitoisuus; työmäärä; luokitukset; RAFAELA-järjestelmä; kuormitus; henkilöstöresurssit; sairausloma; validiteetti; regressioanalyysi; terveydenhuolto

ACKNOWLEDGEMENTS

Everything began more or less unintentionally. A project was started in the first half of the 1990s at Vaasa Central Hospital, which aimed at selecting, developing and implementing a nursing intensity measure and a whole system of patient classification. At first, as an ordinary physician, this activity had no bearing on me whatsoever. However, little by little, I found myself first in discussions about the need for another measure alongside the OPC and other development of the system, and then analyzing one ward's results, followed by those of the whole hospital, then those of all the involved hospitals in Finland. Next, I found myself as a researcher in the Finnhoitoisuus project, under the auspices of the Association of Finnish Regional and Local Authorities, and writing scientific articles on this fascinating topic – and finally, writing my doctoral thesis on it!

I express my warmest thanks to my both supervisors, Professor Juha Kinnunen and Professor Olli-Pekka Ryynänen, for their many excellent comments and intelligent perspectives and their support to me during the various phases of this process. My chief supervisor, Professor Juha Kinnunen has been flexible, tolerant and patient indeed, and has given me all the support, encouragement and, on the other hand, all the freedom that I have needed. He has also accepted that every now and then, I have had to allocate all my disposable time elsewhere; to scientific or other activities.

I am also very grateful to my pre-examiners, Professor Arja Isola and Professor Kari Lindström, for their valuable comments and recommendations, which, in the final phase, improved the quality of my thesis.

Likewise, I appreciate the useful comments that I received from Professor Kaija Saranto, Professor Jari Vuori and statistical comments from Kari Mauranen, M.Sc.

This thesis would never have been started without my enthusiastic, innovative, committed and scientifically-oriented friends, Professor Lisbeth Fagerström – who first suggested to me that I should proceed with this study topic to the doctorate level – and Anna-Kaisa Pusa, Ph.D. Together we three have created the scientific base for the RAFAELA system. It has been a great pleasure to see that "our" system has become widely and ever increasingly used in Finland and is now beginning to spread also to other countries. However, this success – and my thesis – could not have been possible without many other innovative and committed people: Ritva Larjomaa, Development Manager; Toini Nukari, Product Manager; and the person who started the whole project in Vaasa, Chief Nursing Director Kristina Anttila. Warm thanks also to my friend and co-author, Kati Ojaniemi, M.Sc., for all her valuable contributions during this process and to Head Nurse Rose-Marie Swanljung, who has been an important resource with the RAFAELA for years.

I have been happy to have had the opportunity to make the acquaintance of Professor Mika Kivimäki, and I have learned a great deal from him in the field of research. Likewise, I am grateful to my other co-authors in various sub-studies: Professor Jussi Vahtera, D.Med.Sc.; Marianna Virtanen, Ph.D.; Marko Elovainio, Ph.D.; and Kaija Nojonen, Ph.D.

I am grateful to Joanne Jalkanen, M.A., M.A. Eds. for her competent proof-reading and text-editing of the English language, and for her kindness and flexibility. Warm thanks also to my friend, Pia Wik, M.Sc., for translating the abstract into Swedish.

The positive attitude and support of my superiors, Director Göran Honga and Medical Director Timo Keistinen, have promoted my work. Likewise, the good, merry and spontaneous atmosphere at work, and my trustworthy and flexible workmates in the administration of the conservative division and in the hematological ward during these years, Anders Almqvist, Marie-Louis Berglund, Marianne Buss, Åsa-Britt Forth-Snellman, Tarja Holm, Vesa Kataja, Kaj Lahti, Boel Ludén, Carita Mäkelä, Anneli Nyqvist, Kirsi Vaaranmaa and many nurses – too many to mention personally – have made my scientific path easier to tread. The Medical Research Fund of Vaasa Hospital District has also supported my thesis, and for this I am grateful.

Last but not least, the support, flexibility and patience of my own family have been of crucial significance. I am deeply grateful to my daughters Sini, Satu and Taina, and their husbands Martti and Topias, and, of course, to my wife Benita, for all the support that I have received from each of them.

Vaasa, August the 16th, 2008

Auvo Rauhala

LIST OF ABBREVIATIONS

BeNMDS	Belgium Nursing Minimum Data Set
DRG	Diagnosis Related Groups
HRM	Human Resource(s) Management
HSSG	Hospital Systems Study Group
OPC	Oulu Patient Classification. N.B., the uniform national version of the OPC, administered by FCG Efeko Oy, is called the OPCq. In this thesis, however, both of these are called the OPC.
PAONCIL	Professional Assessment of Optimal Nursing Care Intensity Level
PCS	Patient classification system
SHRM	Strategic Human Resource(s) Management

LIST OF ORIGINAL PUBLICATIONS

1. Fagerström L, Rainio A-K, Rauhala A, Nojonen K. 2000. Professional Assessment of Optimal Nursing Care Intensity Level. A New Method for Resource Allocation as an Alternative to Classical Time Studies. *Scandinavian Journal of Caring Sciences 14, 97-104.*

2. Fagerström L, Rainio A-K, Rauhala A, Nojonen K. 2000. Validation of a new method for patient classification, the Oulu Patient Classification. *Journal of Advanced Nursing 31 (2), 481-490.*

3. Rauhala A, Fagerström L. 2004. Determining optimal nursing intensity: the RAFAELA method. *Journal of Advanced Nursing 45 (4), 351-9.*

4. Rauhala A, Fagerström L. 2007. Are nurses' assessments of the workload affected by non-patient factors? An analysis of the RAFAELA system. *Journal of Nursing Management 15 (5), 490-9.*

5. Rauhala A, Kivimäki M, Fagerström L, Elovainio M, Virtanen M, Vahtera J, Rainio A-K, Ojaniemi K, Kinnunen J. 2007. What degree of work overload is likely to cause increased sickness absenteeism among nurses? Evidence from the RAFAELA patient classification system. *Journal of Advanced Nursing 57 (3), 286-95.*

AUTHOR'S CONTRIBUTION TO ARTICLES

Articles 1 and 2: A.R. was responsible for the data analysis, took part in the study design and drafting of the manuscript, and made critical revisions to the manuscript.

Articles 3 and 4: A.R. was responsible for the study design, data collection and analysis and was mainly responsible for the drafting of the manuscript.

Article 5: A.R took part in the study conception and design and data analysis and was mainly responsible for the drafting of the manuscript.

List of tables

List of figures

Contents

1 Introduction

1.1 The need for measuring nurses' workload using a reliable patient classification system

In recent years, the claim has been made that there is an insufficiency of both human and economic resources in health care. However, there is no lack of patients in need of increasingly complicated and sophisticated medical interventions and nursing care. In the search for solutions to this contradiction between the demand and supply of health services, part of the solution lies in the aims of HRM (human resource management): the simultaneous occurrence of high performance, high occupational welfare and job satisfaction. According to the most fundamental idea of HRM, it is the human capacity and its commitment which, in the final analysis, distinguishes successful organizations from the others (Storey 1995).

To overcome scanty personnel resources and to succeed in the recruitment and retention of employees in the increasingly competitive environment of health care, individual organizations, and the whole health care system, must pay more attention to workers' job satisfaction and welfare. Illness and sick leave that may have their origin in excessive workload must be minimized. This dimension of stamina and occupational welfare in HRM is emphasized in this study. Conceptually, however, the definition of the HRM is much wider, as I shall show in detail in Chapter 3.

It is not enough that the productivity and the welfare of personnel are presumed to be adequate. In order to be able to tackle these new and more demanding challenges, the health care sector has been forced to pay more attention to developing information systems that monitor or control various aspects of its activity; processes, outcomes, personnel and economics. Everything must be measured and proved in a creditable and plausible way. Instead of economic or norm control, various comparisons (e.g. productivity and effectiveness benchmarking) between health care organizations have attained an increasingly central position (Fagerström & Rauhala 2003, Junnila 2004, Fagerström & Rauhala 2007).

The largest working group – and cost – of specialized health care is nursing personnel. Therefore, it is important to monitor nurses' workload, its optimality, allocation and costs and to compare the results to corresponding results in other organizations. Similarly, plans and interventions based on the results of measuring and monitoring are a central part of the strategic and operative management of nursing and the entire health care sector. In this task, managers previously had to be satisfied with simplified parameters, such as nurse-to-bed or nurse-to-patient ratio and the leaders' subjective impression and assessment of the situation. However, during the last few decades – and especially in recent years – the situation has changed profoundly. Alongside the development of information technology and nursing workload measurement systems, it has become possible to manufacture user-friendly integrated information systems that can produce a variety of information concerning nurses' workload and the parameters associated with this. It is said that these systems have made the invisible nursing work visible. (Fagerström 1999). Moreover, as the labour market crisis associated with nurses' resignations in the autumn of 2007 in Finland has shown, the functioning of the whole specialized health care is completely dependent on the availability of this nursing workforce.

In recent decades, numerous nurses' workload measurement instruments and patient classification systems have been created (Carr-Hill & Jenkins-Clarke 1995). However, most of them have not been adequately studied and validated scientifically. Measurement systems should be tested for practicability, reliability and validity before being promoted or adopted on a wide scale. Merely making nurses' workload visible is insufficient; the measures must also provide correct and creditable information about it.

1.2 The general presentation of this study

In general terms, this thesis will present, analyze and discuss phenomena in the sector of health care administration, such as measures and their validation, measurement of the workload of nurses, patient classification systems and other nurse demand evaluation methods, staffing and human resources management. This study belongs to the field of health care management sciences.

The RAFAELA system is used as a case, as an example that is presented more closely and thoroughly. Thus, the more specific object of this study is the RAFAELA patient classification system. This is an advanced type of PCS that, alongside measuring nurses' patient-associated workload, also allows a quite unique possibility for the comparison of their workload to an optimum. The rapidly increasing use of this PCS in Finland and the international interest it has aroused caused the need to strengthen and assure the scientific basis of this system.

The RAFAELA patient classification system was created, developed and studied in Finland during the 1990s. Since the national Finnhoitoisuus project, concerning benchmarking in nursing care in 2000 – 2002 (Fagerström & Rauhala 2001, Fagerström & Rauhala 2003), this system has become widely and more increasingly used in Finland. As such, it has become by far the most used PCS in specialized health care in Finland. It is an advanced system of patient classification, which also allows an analysis of workload in relation to an optimum. The workload of nurses can be measured by patient classification systems, based on the measurement of nursing intensity. However, it is important to make the distinction between patient-related and ward-related workload. Nursing intensity measures only the patient-related, direct and indirect, workload of nurses and does not include their ward-related work (Morris *et al*. 2007).

At a more general level, the aims of this thesis are to evaluate whether the RAFAELA system is valid and feasible enough to be used as a measurement tool for HRM in nursing in the wards of Finnish specialized health care. In this sense, this thesis is also of interest and has a role within national health and personnel politics. The standard OPC measure is planned for use in usual hospital wards at specialized health care hospitals. Therefore, the scope of this thesis is also limited only to wards of specialized health care units.

A summary of the study design of this thesis is presented in Figure 1. The objects that have been studied (the shaded objects at the upper edge of the figure) are the measures of the RAFAELA system (the OPC, the PAONCIL and nursing resources) and their more sophisticated combinations (nursing intensity per nurse and optimal nursing intensity), likewise also non-patient factors and sick leave. Non-patient factors mean factors other than nursing intensity that might affect nurses' experience of their workload, for example, administration, co-operation and so forth. The six study questions of the sub-studies (Q1. –

Q6.) are presented in the middle of the figure, using a shortened form of their titles (for the actual titles in full, see Table 8 on p. 80). Their links to the objects are shown with grey arrows. The links of study questions to these sub-studies (at the lower edge) are shown with black arrows. The seventh study question is more comprehensive, analyzing the role of the RAFAELA system in the HRM of nursing. HRM (at the right) was used as one theoretical approach for the study, and the evaluation of the validity and feasibility (at the left) of the RAFAELA system was the central methodological idea of this thesis.

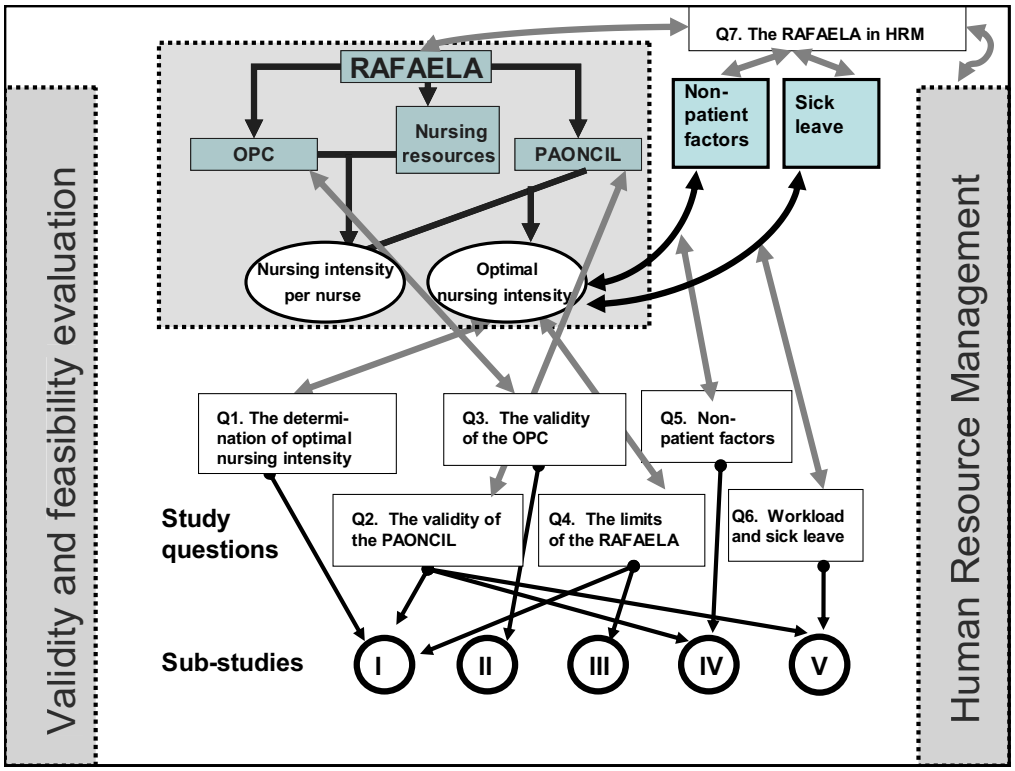

Figure 1. A summary of the study design of this doctoral thesis.

See the text for explanation of the figure.

2 Patient classification systems and other nurse demand evaluation methods

The data used for this thesis and its five sub-studies comprises research articles and, to a lesser degree, review articles and books. For the most part, PubMed and Cinahl-databases were searched, using various combinations of keywords: workload, workload measurement, personnel staffing and scheduling, health workforce planning, classification, patient classification, nursing administration, staffing level, nursing (care) intensity, patient acuity, instrument validation, validity, reliability, stress, occupational, sickness absence, human resource(s). This more systematic part of literature was completed using direct searching from journals – Journal of Advanced Nursing; Journal of Nursing Management; Human Resource Management; and Journal of Management. Moreover, the literature gathered in the HumanRe project of the Department of Health Policy and Management of Kuopio University was available. Finally, articles were received from co-authors and other researchers in the field of health care management science.

2.1 Increasing demands for estimating adequate nursing staffing

For a long time, the performance of whole hospitals has been observed using crude, overly simplistic and biased measures: inpatients' bed-days and hospital care periods, outpatients' visits, patients' number of operations, and so on. These statistics are prone to many kinds of errors, misinterpretations, bias and manipulation. Therefore, they have not generated sufficient systematic information that could have been used to support clinical or administrative decisions (McCallum 2004). The quality of care and the need for resources, based on the complexity of cases (case mix) of hospital services, have not been sufficiently taken into consideration in these statistics. In recent years, the increasing use of DRGs (Diagnosis Related Groups) has made it possible to supply the above-mentioned parameters with weight coefficients that describe better their degree of difficulty or complexity and cost consumption. Thus, at national, hospital, specialty, ward and patient level, it has become possible to make more comprehensive performance analyses and to do productivity benchmarking in health care by using a DRG system and its weight coefficients (Lauharanta *et al*. 1997, Junnila 2004).

In the same way, nursing staffing and estimation of nurses' workload have traditionally been based on similar crude and oversimplified data, which have proved inadequate for the formation of a sufficient information basis for HRM in nursing. Initially, the complexity of the subject of staffing was not taken sufficiently into consideration (Arthur & James 1994). For a long time, the main interest has been directed at parameters such as nurse-to-bed and nurse-to-patient ratio, nursing hours per patient day, and so forth. Influenced by the scientific management approach to work analysis, the emphasis lay in analyzing nursing tasks and the time taken to complete them, calculating the whole-time equivalent the nurses needed to complete a specified volume of tasks (Auld 1976). Patient dependency measures have provided another, more enduring approach (Adams & Bond 2003). Efforts to predict nursing workload from case mix groups based on diagnostic codes and procedures (e.g., DRG or Diagnosis related Groups and the British HRG or Health Resource Groups) have failed. In the study of Campbell *et al*. (1997), for example, the case mix group accounted for only 18% of the variation in nursing time required.

In the European Union, the number of hospital beds has decreased. Eurostat data (Internet reference No 1: European Commission Health & consumer protection directorate-general) show that the total number of hospital beds has decreased substantially in most Member States since 1980. For EU 15, the number of hospital beds fell by more than 30% between 1980 and 2000. A considerable part of this reduction is likely to have been caused by the drop in the length of hospital stay. This decreased in EU 15 from 17.4 days in 1980 to less than 11 days in 1997. Another explanation is that an increasing number of patients are treated on an outpatient basis, and only the sickest are treated as inpatients. The increased role of day case surgery plays an important role in operative specialties.

As reported above, the length of stay in hospital becomes increasingly shorter. Nowadays, only the sickest and eldest patients with major comorbidity are treated as inpatients. Ever more advanced and sophisticated technologies are used in health care. As a result of advances in supportive therapies, it is possible to treat patients who earlier were considered as beyond all therapeutic interventions. People's demands have also increased. Consequently, inpatients in specialized health care have become more care-intensive and demanding in recent years.

The amount of nurses per population has not increased notably in recent years. According to European Union statistics (Internet reference No 2: European Commission Health & consumer protection directorate-general), the number of qualified nurses and midwives per 100 000 inhabitants changed between the years 1995 and 2001 in the EU 25 and EU 15-index, correspondingly, from 734 to 748 and from 752 to 753 nurses and midwives per 100 000 inhabitants. Despite this slight growth in the reported number of nurses, in many countries, the availability of nurses is insufficient to meet all of the health care demands, and it is anticipated that this situation will worsen in coming years (Buerhaus *et al*. 2000, Janiszewski Goodin 2003).

In an Australian study (Hegney *et al*. 2003), 30 – 32% of nurses in the acute sector reported difficulties in meeting patient needs because of insufficient staffing levels. A study (Aiken *et al*. 2002) concerning 43 000 nurses in five Western countries showed that 17 – 41% of nurses were dissatisfied with their present job. They also found that 27 – 54% of nurses younger than 30 years of age planned to leave their organization within the next year. A central cause for dissatisfaction was insufficient staffing levels. In addition, their job dissatisfaction and emotional exhaustion (burnout) was increasing. In her thesis (2005), Marjukka Laine studied the commitment of Finnish nurses to their job. Every fifth nurse has been considering changing his/her job and every sixth nurse has been considering leaving his/her profession completely. According to her study, most Finnish nurses are strongly committed both to their workplace and profession. The factors associated with a rewarding job had a greater significance for commitment than all the requirements of the job. Nurses' commitment was especially decreased by their experiences that their own work was not considered as important, their possibilities to develop or influence were weak, their capabilities were underused, their workplace had a poor atmosphere, they experienced their management as poor, and they considered their workload as excessive. Thus, Laine's thesis also emphasizes the significance of workload optimality as a central aim of HRM in nursing.

There are great differences between wards in the case mix of patients and in the skill mix of nurses (the proportion of registered or professional nurses, second level or licensed practical nurses and unlicensed personnel or nursing assistants) and in the processes, quality and outcomes of nursing care. (Sermeus *et al*. 2007). A critical mass of literature has established that higher nurse staffing levels in North American hospitals are associated with decreases in adverse patient events and negative outcomes for nurses. Studies have generally shown that

better staffing is associated with lower risk-adjusted hospital mortality (Aiken *et al*. 2002, Needleman *et al*. 2002, Sermeus *et al*. 2007). Aiken *et al*. (2002) found, for example, that every unit increase in average patient per nurse ratio in adult general hospitals produced a 7% increase in the risk of mortality and failure to rescue after controls. This equates to five fewer patients per 1000 of the types studied being expected to die in hospitals where ratios were 4:1 versus 8:1. Multiple studies show an association between nurse staffing and a variety of non-mortality outcomes, but findings with respect to particular indicators have not been consistent. A recent systematic review of North American research confirmed this by concluding that increased nursing staffing in hospitals is associated with lower hospital-related mortality, failure to rescue, and other patient outcomes (Kane *et al*. 2007). However, Kane *et al*. (2007) concluded that the nursing effect must be conditioned by provider characteristics including hospital commitment to high quality care, not considered in most of the studies.

As a conclusion, the decreased number of hospital beds is occupied by increasingly old and sick patients with greater comorbidity, who place greater demands on the quality of nursing and on the exploitation of more advanced medical technology. At the same time, the amount of nurses has not increased markedly. Therefore, nurses report difficulties in meeting patient needs because of insufficient staffing levels. Their job dissatisfaction and emotional exhaustion is increasing. Because of this, a significant proportion of nurses plan to leave their organization. It is expected that this shortage of nurses will increase in the coming decade because of increased demand related to the aging of populations. Moreover, the detrimental effects of nurses' understaffing and excess workload to patient outcomes have been scientifically established. Thus, an optimal and just allocation of the limited nurse resources is, in many ways, of crucial importance. In the HRM of nurses, all these factors together have caused an increased need to measure nurses' workload and to estimate adequate staffing using a valid and feasible measurement tool.

2.2 Concepts, definitions

Nursing intensity

Nursing intensity means the assessment of the amount and competence of nursing care that patients need, especially for adequate staffing and/or for calculation of costs (Partanen 2002).

In the literature, concepts such as nursing intensity and patient classification (see below) are sometimes used as synonyms. Nursing intensity is also referred to as '(patient) acuity', 'severity', 'care intensity', 'intensity of care', 'nurse dependency', 'patient dependency' and 'nursing care intensity' (Fagerström 1999, Partanen 2002). The level of direct and indirect patient-related nursing work activities is gauged by way of nursing intensity (Morris *et al*. 2007).

Patient classification

Patient classification is a generic term that denotes any grouping of patients, for example, according to diagnosis, treatment, DRG, blood group, demographic factors, and so on. Thus, the term patient classification can also be used when patients are grouped according to their nursing intensity. In this narrow meaning, it is defined by Giovannetti (1979) as a categorization of patients according to some assessment of their nursing care requirements over a specified period of time. The primary aim of patient classification is to be able to respond to the constant variation in patients' caring needs (Huckabay 1981).

Patient classification system

Patient classification system is also a wide term that, in association with nursing intensity, refers to the identification and classification of patients into care groups or categories, and to the quantification of these categories as a measure of the nursing effort required (Giovannetti 1979). According to Shaha (1995), patient classification systems, also known as patient acuity systems, have been developed to manage workloads by identifying and quantifying patient care needs to estimate the required staff for each patient (Strickland & Neely 1995).

Many of them are produced for general use, but most of them are planned for a specific patient group and/or environment, such as a dialysis unit, intensive care unit, rehabilitation ward, neonatal intensive care unit, and so forth.

Nursing workload

No common definition of nursing workload exists in the literature. According to the proposal of Morris *et al*. (2007), nursing workload should represent the sum of nursing intensity (direct and indirect patient care activity) and nursing activities not related to patient care.

The term 'nursing workload measurement' refers to any attempt to assess the volume and/or level of nursing work. One of the functions of workload measurement systems is to determine nurse staffing requirements, but this is not their sole purpose (Arthur & James 1994).

Thus, the terms 'nursing intensity' and 'nursing workload measurement' are not identical. According to Partanen (2002), 76% of registered nurses' and licensed practical nurses' working time was directed at the patients' direct and indirect care activities. Unit-related work activities accounted for on average 16%, and the time distributed to personal activities eight per cent. Nursing intensity measures only the patient-related, direct and indirect, workload of nurses and does not include their ward-related work (Morris *et al*. 2007). Therefore, it is important to make a clear theoretical distinction between the concepts of nursing intensity and workload.

Nurse demand methods

A nurse demand method refers to any system of determining the number and/or mix of nursing staff. A nurse demand method may (but does not necessarily) use some instrument of workload measurement (Arthur & James 1994).

2.3 A taxonomy of nurse demand methods

There are numerous nurse demand methods and nursing workload measurement systems – and different ways to classify these systems, some of which are defined as PCSs. According to Arthur and James's (1994) overview of methods for determining nurse staffing requirements, a taxonomy of nurse demand methods consists of six different systems (Table 1).

Consensus approaches are overtly subjective and attempt to take a critical and reflective view of nursing workload. Management approaches aim at being objective and prescriptive. 'Top-down' and 'bottom-up' approaches differ in the level at which information is obtained for the calculation of staffing requirements, by using ('top-down') national or regional recommendations or statistical formulae, or ('bottom-up') local, ward-level information.

All the methods have their problems and limitations. The intuitive method is the oldest known method. Its main problem is its subjectivity and the lack of consistency across different wards/units and between hospitals that makes comparisons difficult. The consultative methods suffer from the same problems and limitations. Staffing norms and formulae have the limitation of being insensitive to local variation.

Table 1. A taxonomy of six nurse demand methods according to Arthur and James, 1994

Approach	**Type of method**	**Description**
Consensus approach	1. Intuitive method (descriptive method; professional judgement)	Assessment of staffing requirements by nurse-in-charge from all available information.
	2. Consultative method	A structured review of a ward's operations and patient activity and an assessment of required nurse establishment.
Top-down management approach	3. Staffing norms	Recommendations by professional bodies; may take the form of nurse-to-bed ratios.
	4. Staffing formulae	A statistical relationship between a set of independent variables (e.g. through-put, beds, demography of locality) and nursing workforce.
Bottom-up management approach	5. Nursing intervention	Measures time of nursing interventions, procedures and tasks.
	6. Patient dependency	Categorizes patients into groups according to dependency. A workload index is calculated from the sum of weighted categories.

Nursing intervention methods involve an assessment of the patient for the activities involved in producing the nursing care required. Through observation, each activity has been allocated a time to carry it out. The sum of a patient's required nursing activity times should supply the total time required to care for that patient in that shift or day (Hughes 1999). These methods tend to be task-orientated; they tend to prioritize physical nursing and overemphasize the mechanistic process of nursing. Another problem is that nursing is seen as a linear activity. This ignores nurses' ability for 'multitasking'; to meet several needs simultaneously. Likewise, it may fail to make adjustments for skill mix (Hunt 1990).

Patient-dependency methods are based on the belief that the dependency of the patient on a nurse is a good measure of the demand on nurses' time. Patients are placed into categories according to their level of 'dependency'. Each category is weighted and thus, a workload index can be calculated. Since they do not require as detailed information as an intervention method, dependency methods require a simpler assessment.

According to Carr-Hill and Jenkins-Clarke (1995), approaches to measuring nursing workload can be classified into four groups (Table 2).

Table 2. Four approaches to measuring nursing workload (Carr-Hill & Jenkins-Clarke 1995)

Approach/method	**Description**
1. Dependency-driven	Workload requirements are based mainly on the dependency of ward patients on a certain amount of nursing care in order to perform the basic activities of daily living.
2. Task-oriented	Based on the recording and predicting of nursing interventions for individual patients.
3. Care plan driven	Workload is measured by producing nursing care plans which are then used to predict workload.
4. Ward-based	Ward overviews of staffing requirements are produced by concentrating on patient through-put/bed occupancy.

The methods 1, 2 and 4 of Carr-Hill and Jenkins-Clarke (1995) correspond to the methods 6, 5 and 4 of Arthur and James (1994). The methods described are not mutually exclusive. According to Arthur and James's (1994) classification, for example, many PCSs belong both to nursing intervention and to patient-dependency methods.

In the RAFAELA system, the OPC can be considered as an advanced form of a 'bottom-up' management approach and mainly a patient-dependency method without separate categories. However, the OPC measures only such needs that have been responded to and, therefore, here it behaves like a task-oriented/nursing intervention method. The PAONCIL can best be

considered a kind of bottom-up method that is intuitive/descriptive/professional, with a structured scale with definitions and descriptions.

Another approach in assessing nursing resources is the concept of nurse dose. It is conceptualized as having three equally essential components: dose, nurse and host and host response. Dose, in the macro view, includes the number of nurses per patient or per population in cities, countries, and so on. In a micro view, dose includes the amount of nurse time and the number of contacts. The nurse component comprises the education, expertise and experience of the nurse. Host is represented by an organization and its characteristics (culture, autonomy, practice control) in a macro view, and by the patient and characteristics (beliefs, values, culture) in a micro view. Host response includes response to the autonomy and acceptability of the nurse. According to this approach, optimal patient outcomes and health care cost savings can only be achieved with the appropriate nurse dose, with its three equally essential components (Brooten & Yongblut 2006).

2.4 The creation of patient classification systems

2.4.1 The general process of constructing a new instrument

The building of an instrument can emerge from a practical need or from theoretical considerations. In both situations, however, the building of an instrument begins with familiarization with the underlying theories on the field of the phenomenon that is of interest. The next phase is to ask or find the relevant question for which an answer is sought. Then, the creator of an instrument must choose one of the many existing theories as a basis for the construction of the instrument. Following this, these theoretical terms must be operationalized, in other words, a definition of them must be found that is measurable. After this, the measurement instrument is initially constructed from separate items. The items are then critically examined, validated and edited. The instrument is then tested in a pilot study and, based on this, it is again corrected and changed, if necessary. Only after this can the measure be considered as ready for wider use (Karma 1987, Streiner & Norman 1995, Metsämuuronen 2003, Sermeus *et al*. 2004).

2.4.2 The choice of critical indicators

Perhaps the most common error in creating new measures is to dismiss existing measures too lightly. In the exhausting array of available measures, there may still exist one that is suitable – or that can be made suitable with minor changes (Streiner & Norman 1995). If the decision for the creation of a completely new measure is made, then the next question encountered concerns which of the two main types of patient classification systems to choose: prototype and factor-evaluation system. Characteristic for prototype classifications were extensive descriptions of typical patients (prototypes) or typical nursing tasks in every patient class. Prototype classifications were more popular at the beginning of the era of PCSs, but have been replaced by factor-evaluation systems. They are characterized by several critical indicators or factors of nursing care. These critical indicators are then scored independently of each other. Nursing intensity or patient class is then determined according to the system's algorithm (Giovannetti 1986, Fagerström 1999, Partanen 2002).

The creation of factor-evaluation-type instruments for patient classification is based on underlying assumptions concerning human beings' need for care and caring (Kaustinen 1995, Fagerström & Bergbom Engberg 1998, Fagerström 1999), and often starts from a pre-established standard of care (Giovannetti 1979). In the literature, there are several examples of instruments in which the obvious focus is on the physical needs of the patient and where psychological needs, such as emotional support and need for information and guidance, are lacking; something that has been criticized by several researchers (Giovannetti 1979, Kaustinen 1995).

In creating a nursing intensity measure, the choice of critical indicators of nursing care (descriptors of patients' nursing care requirements), indicates underlying theoretical assumptions. Giovannetti (1979) emphasizes that these critical indicators are supposed to represent the nursing care activities which most decisively influence nursing care time, and at the same time, she stresses that patient classification does not measure a patient's real needs, although it implies a rough estimate of the patient's caring needs and the nursing care time required to satisfy them.

2.4.3 From selection of items to building an instrument

When the critical items have been selected then the scale must be selected. Either a continuous or a type of categorical scale is selected. Categories must be planned and described in such a way that they all gain cases. The scales and their descriptions must also be planned so that biased responses are minimized. Otherwise the instrument loses its efficiency.

The next step is then the weighting of the items, either on a theoretical or on an empirical basis. The empirical approach is done with multiple regression analysis, aiming at selecting coefficients so that the predictive accuracy of the equation is maximized. Another way is to make direct measurements, e.g. time studies, to be able to determine the optimal coefficients or weights (Streiner & Norman 1995). When all this has been completed, lastly, the validity and reliability of the instrument must be examined and secured.

2.4.4 Time studies and other methods for estimating the need for personnel

In addition to an instrument for assessing nursing intensity, systems of patient classification often comprise some form of time study, such as work sampling and time-and-motion study (Kirk 1990, Hendrickson *et al*. 1990, Cardonna *et al*. 1997). On the basis of time study, the hours of work required to meet the patients' caring needs in different categories are calculated. These methods often include a calculation of the average number of nursing hours required for each category or a calculation of standard number of nursing hours for different nursing activities or a combination of these two (Giovannetti 1979*,* Vanputte *et al*. 1985).

Traditional time studies, which are assumed to have been developed on the basis of operational research originating in industry, have been criticized (Jennings *et al*. 1989, Procter & Hunt 1994, Milne & McWilliam 1996). Time studies cannot in a comprehensive way capture the nature of nursing care, which is complex and multidimensional, and is carried out in an individual and personal manner, nor can such studies take into account that nurses often perform several nursing activities simultaneously (Williams 1977, Reitz 1985, Zembala 1993, Noyes 1994). Quantitative aspects of nursing care have been emphasized and the qualitative aspects have not been sufficiently taken into account in traditional time studies. Time

allocation should also be estimated on the basis of results (Finnigan *et al*. 1993). It is also assumed that time studies are insensitive to the constant variation in patients' caring needs and thus to variation in the need for staff (Huckabay & Skonieczny 1981). Time studies are also considered to be both expensive and time-consuming (Hoffman 1988). According to Barnum (1990), thousands of hours of research have been invested in the further development of 'staffing systems' for nursing care, but the benefit and the final results have often proved insignificant. She calls for simple and easily applicable methods that involve workable accuracy, but not meaningless perfection.

2.5. A history of patient classification systems

The need for PCSs was contributed to by early clinical and scientific observations that patients had a varying need for nursing care, both quantitatively and qualitatively, and both between different patients and in the same patients between different days (Partanen 2002). The primary purpose of patient classification has been to match perceived patient needs with available nursing resources (Alward 1983, Shaha 1995, Strickland & Neely 1995). Systems for patient classification have been considered to provide rational, systematic and objective bases for staffing and budgeting of staffing expenses (Giovannetti 1979, Nagaprasanna 1988, Kelleher 1992, Levenstam & Bergbom Engberg 1993, Strickland & Neely 1995).

Systems for patient classification have been developed predominantly in the United States of America (USA) since the 1950s, but also in the United Kingdom and also, for example, in the Nordic countries since the beginning of the 1960s (Giovannetti 1986, SPRI 1990, Hlusko & Nichols 1996). A multitude of patient classification systems exist. Already in 1973, Aydelotte reported 40 types of patient classifications. In the 1980s, most (42%) hospitals in the USA used an internally-developed patient classification system and only 16% used commercialized systems (Nagaprasanna 1988).

As was already mentioned in Chapter 2.4.2, the development of methods of measuring has led to two main types of instrument for the classification of nursing intensity: prototype and factor-evaluation systems. Most of the earlier PCs belonged to the prototype group. Factor-evaluation systems were in wider use in the 1990s (Seppälä 1992). They are characterized by

several critical indicators or factors of nursing care. These are then scored independently of each other and nursing intensity or patient class is then determined according to the system's algorithm (Giovannetti 1986, Fagerström 1999, Partanen 2002). Internationally, the most famous factor-evaluation systems have been the Rush-Medicus system, the Grasp system, the Public Health Service Patients Classification System and the Canadian Hospital System Study Group (HSSG) patient classification (Levenstam & Bergbom Engberg 1993). The HSSG system has also been the basis, for example, for the Oulu Patient Classification (OPC) and Monitor systems (Partanen 2002), and, via the OPC, also partly the basis for the RAFAELA system.

Traditional systems for patient classification, developed since the 1950s in the USA, measured the nursing intensity of patients. The need for personnel resources was then calculated on the basis of time studies and activity analysis (Giovannetti 1979, Alward 1983, Rosenbaum *et al*. 1988, Giovannetti & Moore Johnson 1990, Phillips *et al*. 1992, Mayo & Van Slyck 1999). According to Arthur and James (1994), these nursing intervention methods were relatively expensive to implement and were also developed from a task-oriented approach, thus, physical nursing care was prioritized over non-physical care.

There have been three generations of PCSs in the USA. Their more widespread use began there in the early 1960s. Estimated annual staffing needs were calculated manually with first generation PCSs. A basic for prediction was formed by historic data of patients per nurse (Malloch & Conovaloff 1999).

In the 1980s, the introduction of managed care and DRGs caused deep changes and new demands in US health care. High productivity and efficient resource allocations became increasingly important in order to make optimal use of resources (Hagerty & Spengler 1985, Rieder & Lensing 1987, Giovannetti & Moore Johnson 1990). The development in technology and informatics also assisted in developing the second generation PCSs in the 1980s. The focus of these was the monthly staffing budget. The first and second generation PCSs were both criticized for three reasons. They lacked credibility because of representing a 'robotic' industrial model. They failed to notice multitasking of personnel and they also failed to detect the variability of caregivers.

The 1990s saw new demands in the battle between downsizing the workforce and legislating mandated minimal staff levels. Clinical outcome measures were claimed for accreditation. According to Malloch and Conovaloff (1999), third generation patient classification systems were developed to meet these new requirements and in response to the above-mentioned criticism towards earlier PCSs. These systems were focused on optimizing the patient-caregiver process for the outcomes of cost, health and caregiver satisfaction. Even in the United Kingdom and Nordic countries, patient classification was an ongoing topic in the literature during the 1980s and 1990s. The aim of many projects has been to find a reliable and useful tool for nursing management and management of personnel resources (SPRI 1990, Levenstam & Bergbom Engberg 1993, Arthur & James 1994, Carr-Hill & Jenkins-Clarke 1995, Campbell *et al*. 1997, Hughes 1999).

A patient classification system project was started in 1991 at Oulu University Central Hospital in Finland. The aim was to develop a new nursing intensity measure because available instruments were not considered suitable for use in the context of Finnish specialized health care. The HSSG (Hospital Systems Study Group) PCS (Goldstone *et al*. 1985) of Canadian origin was selected as the basis for the development project. It was considered that this original PCS overemphasized the patient's physical needs and responding to them. This project led to the development of the OPC (Oulu Patient Classification) instrument (Kaustinen 1995). The OPC instrument was then further developed at Vaasa, and another completely new measure, the PAONCIL was also developed. Thus, by combining these two and information about daily nursing resources, the RAFAELA patient classification system was born. For more information on the need for a new PCS, the RAFAELA system, its history and development, see Chapter 5.

2.6 The present use of patient classification systems

At present, there are numerous patient classification systems in various parts of the world. However, many of them are business secrets and, therefore, difficult to evaluate, especially in the competitive US marketing environment, which is protective of proprietary products (Malloch & Conovaloff 1999). The trend in PCSs is towards the so-called third generation PCSs, which combine nursing intensity information with other important information sources and follow-up parameters to create a more comprehensive information system, supporting

adequate and optimal administrative decision-making (Malloch & Conovaloff 1999). In principle, this can be made by an 'all-in-one', monolithic PCS program or by other programs that integrate data from different sources.

Originally, patient classification systems were planned as an objective basis for staffing. Numerous other uses of patient classification system information have been described in the literature, for example, assessment of patient care needs, budgeting, costing out nursing services, billing, cost control, quality assurance and research. However, by far the most common use of PCS information has been to assist the nursing department with staffing decisions (Botter 2000). In these staffing decisions, even modern PCSs will only allow longitudinal data to be collected concerning nursing workload, to be used to estimate more accurately future labour resources and budgets. However, PCSs cannot accurately predict staffing for the next shift. This failure to meet staff expectations is understandable, because PCSs can neither control nor predict patient admissions, nor can they predict the changes in patient conditions that will change nurses' workload (Seago 2002).

Nowadays, especially in the USA, economic aspects, such as nursing costs, billing and reimbursement have attracted an increasing amount of attention. However, in the USA, direct and indirect nursing expenditures are still rolled up into cost centres and treated as a fixed cost, then billed at daily room rates. Reimbursement for hospital care is based primarily on the medical diagnosis, DRG and principal procedures. Nowadays, hospitals in the USA are not yet reimbursed for different levels of nursing intensity within these DRGs, essentially hiding the variability of nursing care. It has, however, been shown that nursing intensity and estimated direct nursing costs vary significantly within and across similar nursing units, despite being billed at the same daily room and board rate. It is proposed that a new cost centre for nursing must be created so that the DRG is adjusted by daily nursing intensity to allow independent costing, billing and reimbursement of inpatients nursing care (Welton *et al*. 2006). The situation with the DRG costing and billing is similar in Finland. Nursing expenditures are divided evenly between various patients and DRG groups and are not based on the consumption of nursing resources per patient and per DRG group, measured with a PCS (Aaltonen *et al*. 2007).

In California, for example, all acute care hospitals are required to have a reliable and valid patient classification system. The two general types most commonly used there are the

summative task type PCS and the critical incident or criterion type PCS. There is little to assist nurse executives in deciding what type of system to choose. There is only modest research demonstrating the validity and reliability of different types of PCSs and very little published data comparing the predictive validity of the different types of systems available in the US (Seago 2002).

There are several patient classification systems in use in all the Nordic countries, for example, Beakta, Rush-Medicus, Monitor, Zebra and TIC. In Finland, the most-used nursing intensity measure is the Oulu Patient Classification (OPC), either alone or, after some modifications, as part of the RAFAELA system. The Monitor system is in use in Kuopio University Hospital (Savolainen & Töyry 1995). Many systems have something in common; for example, the Zebra and the Monitor systems and the OPC are based on the Canadian HSSG patient classification system.

The Zebra system (Levenstam & Bergbom Engberg 1993, Levenstam & Bergbom Engberg 2002) is widely used in Sweden. Its patient classification instrument belongs to the class of dependency-driven factor-evaluation systems. The Zebra system comprises four parts: (1) the patient classification; (2) the activity study as a validation system (Lake 1982); (3) the staffing situation; and (4) the quality of nursing care (deficiencies in the quality of nursing care, when understaffed). The patient classification measures the direct nursing care activities given to each patient per 24 hours. It has the following components: hygiene; nutrition; observation; mobilization; uncontrolled output and extra need of care. Each of these has one to three determinators (A, B and C), reflecting the dependency level and the level of care given. For each patient a determinator must be checked for hygiene, observation and mobilization. The last two components are only used on occasions when there is an extreme need for nursing care. Each combination of determinators is referred to one of the four categories of direct nursing care. Total time of nursing care per patient and category of care are then produced using periodical time studies and questionnaires. Thus, workload is expressed as an ordinal scale variable (one of the four categories and the average time for that category), not as a continuous quantitative variable, which can be considered as a weakness in the system.

In principle, **The Monitor system** includes the same HSSG patient classification instrument as the Zebra system. In the Finnish version, to assure that the translation was correct, the manual was translated into Finnish and then back into English. Some changes were made to

the instrument for cross-cultural adaptation. The work sampling method was used in Kuopio in quantifying the nurses' working time needed for the four categories (Partanen 2002). In Turku University Hospital in 1994 – 1998, Pulkkinen (2000) developed a new nursing intensity measure, called the NETA (Need to Action) model. This also belongs to the class of dependency-driven factor-evaluation systems. The instrument is not in use at present.

It is a common experience that modern patient classification systems are complex systems that require considerable resources. Implementing any PCS is costly and time consuming. Daily data collection takes time, especially with nursing intervention methods. Maintaining the data system, training and motivating nurses and testing the validity and reliability of the system also require staff resources.

2.7 Summary of the patient classification systems and other nursing demand evaluation methods

In recent years, nurses' staffing and their estimation of workload has become a more important topic because of the lack of nurses and their dissatisfaction with their jobs. It has also been shown that higher nurse staffing levels in hospitals are also associated with decreases in adverse patient events and negative outcomes for nurses. Nurses' staffing and workload estimation has traditionally been based on crude and oversimplified data. Likewise, general systems for patient grouping, such as a DRG, have failed to estimate nursing workload. Thus, specific PCSs or other nurse demand methods are needed for the staffing and workload measurement of nurses. In addition to this main utilization, numerous other uses of patient classification system information have also been described: assessment of patient care needs; budgeting; costing out nursing services; billing; cost control; quality assurance; and research.

The creation of a new scientifically-based measurement instrument is a long and formal process and requires a great deal of criticism. In creating a nursing intensity measure, the choice of critical indicators of nursing care, based on underlying theoretical assumptions of patients' nursing care requirements, is a central task. Scales and weights of items are then planned. In addition to an instrument for assessing nursing intensity, systems of patient classification often comprise some form of time study. Based on this, the hours of work

required to meet the patients' caring needs, for example, in different categories, are calculated. However, time studies have received much criticism.

There are numerous PCSs and other nursing demand methods. Nursing intervention/task-orientated and patient dependency-based methods are the two most usual methods of nursing demand evaluation. The former is based on measurements of time needed for nursing interventions, procedures and tasks. The latter categorize patients into groups according to dependency and a workload index is then calculated from the sum of weighted categories. Systems for patient classification have been developed predominantly since the 1950s, first in the USA. There have been three generations of patient classification systems in the USA. Third generation PCSs were focused on optimizing the patient-caregiver process for the outcomes of cost, health and caregiver satisfaction. There are several PCSs in use in all the Nordic countries. In Finland, the most-used nursing intensity measure is the Oulu Patient Classification (OPC), either alone or, after some modifications, as part of the RAFAELA system. The Monitor system is in use in Kuopio University Hospital. Many systems have something in common.

3 The patient classification system as a part of human resources management in health care

3.1 Definitions and history of the HRM concept

According to Viitala (2006b), personnel management can be divided into three sections that fuse together in practical management and cannot be unambiguously demarcated from each other: 1) Leadership; 2) Human Resource Management; and 3) Industrial Relations. Human resource management (HRM) is a distinctive approach to employment management which seeks to achieve competitive advantage through the strategic deployment of a highly committed and capable workforce, using an integrated array of cultural, structural and personnel techniques (Storey 1995). HRM is both a scientific concept and a business practice.

Conceptually, one central aspect of strategic human resource management (SHRM) is the integration of human resource function with strategic decision making (Guest 1989). The basic idea is that the competitive strategy selected can be executed only by means of sufficient, competent and committed personnel (Becker & Huselid 2006, Viitala 2006b). Therefore, personnel strategy must specify what the personnel resources are that the organization needs to reach its goals. These goals are, in turn, specified in its business strategy (Sydänmaanlakka 2004). In this context, patient classification systems that can aid in decisions on necessary staffing resources are a rather relevant part of HRM. Nowadays, these two concepts – SHRM and HRM – are, to some degree, used interchangeably, because the strategic element is considered such a central part of HRM.

The concept of HRM originated in the United States in the discussions of the 1960s and 1970s (Brewster 1993). Raymond Miles (1975 in Sädevirta 2004) was the first systematic HRM theorist. He considered that of the three management theories, the human resources model was superior to the traditional model and human relations model (Ollila 2006, Viitala 2006b). According to his human resources model, an organization consisted of organizational and human variables. The main task of managers was to integrate these variables into an efficient socio-technical system. (Sädevirta 2004). Personnel were no more primarily considered a variable cost to be minimized, but a resource for achieving the organization's aims (Storey

1989). Patient classification systems support this kind of HRM thinking, when also making the nursing work visible, not only its cost.

3.2 Essential features of HRM

The most fundamental idea in HRM is that, among all the factors of production, it is the human resource which really makes the difference. It is the human capacity and commitment which, in the final analysis, distinguishes successful organizations from the others (Storey 1995). In addition, according to HRM, the labour resource should be utilized to its full capacity (Storey 1989). Essential features of HRM (Storey 1989, Vuori 2005) are presented in Table 3. According to HRM, competitive advantage is reached through the workforce; hence all efforts are made to assure that a sufficient, capable and committed workforce is available. Therefore, recruitment, performance, individualism, training, flexibility and organizational climate are emphasized.

Table 3. List of essential features of HRM (Storey 1989, Vuori 2005)

Essential features of HRM
• recruitment is important • performance, not hierarchical position, determines the reward • individual work contracts • division of labour is based on teams • solutions to conflicts between individuals are sought from the organizational climate and culture • working conditions are in harmony with the worker's individual needs • promoting learning and training is central • flexibility is emphasized in following rules • the importance of measuring

According to Storey (1995), the key elements of HRM can also be presented as in Table 4. The data in Tables 4 and 5 is used in Chapter 8.3 as a basis for a list of the central principles of HRM to evaluate the theoretical associations of HRM and the RAFAELA system.

Table 4. The HRM model according to Storey (1995)

1. Beliefs and assumptions • It is the human resource which gives the competitive edge. • The main aim should not be mere compliance with rules, but employee commitment. • Therefore employees should be very carefully selected and developed.
2. Strategic qualities • Because of the above factors, HRM decisions are of strategic importance. • Top management involvement is necessary. • Human resource policies should be integrated into the business strategy – stemming from it and even contributing to it.
3. Critical role of managers • Because human resource practice is critical to the core activities of the business, it is too important to be left to personnel specialists alone. • Line managers need to be closely involved both as delivers and drivers of the human resource policies. • Much greater attention is paid to the management of managers themselves.
4. Key levels • Management culture is more important than managing procedures and systems. • Integrated action on selection, communication, training, reward and development. • Restructuring and job redesign to allow devolved responsibility and empowerment.

The meaning of the concept of HRM is, however, said to be quite ‘elastic’. At its most basic, HRM represents a set of managerial initiatives. It is also said to be a near-equivalent of the concept of ‘the management of change’. It can be used in a restricted sense for only that approach to labour management which treats labour as a value asset rather than a variable cost and which, therefore, recommends investing in the labour resource through training and

development and through measures designed to attract and retain a committed workforce. On the other hand, the concept of HRM can be used to refer to any modern managerial initiatives related to human resources. There is also a 'hard' and 'soft' version of HRM. The 'hard' one emphasizes the calculative and business-strategic aspects of managing the employees in as rational a way as any other economic factor. The 'soft' one traces its roots to the human-relations school and emphasizes communication, motivation and leadership (Storey 1989).

According to Legge (1989), at the normative level (i.e. how practice should be) there is little difference between HRM and personnel management. However, HRM focuses more on what is done to managers, whereas personnel management appears to be something performed on subordinates. HRM also casts line managers in a more proactive role and HRM emphasizes top management's responsibility for managing culture.

3.2.1 Balance in skill and competence, staff mix; human resource planning

As Storey (1995) has expressed, it is the human resource that gives the competitive edge. Personnel are the most central quality factor of an organization. Therefore, employees should be very carefully selected and developed. The organization's personnel strategy, recruitment strategy and selection system should support the overall organization strategy. Thus, personnel strategy states the volume of personnel needed. It also tells the distribution of competence (skill mix) needed.

Human resource planning is an essential feature in the ideal type of HRM. It differs from traditional manpower planning through a 'developmental' approach to employees (systematic management of the assessment and augmentation of their ability, in relation to, for example, business needs) and through an emphasis on commitment to the goals of the organization (Rothwell 1995). The main parts or interventions of HRM are the following four key elements of the HRM cycle: selection, appraisal, development and rewards (Figure 2). These procedures should be tackled seriously in order to advance the performance of the organization (Storey 1989). In the HRM of nurses, their performance can be measured with a patient classification system.

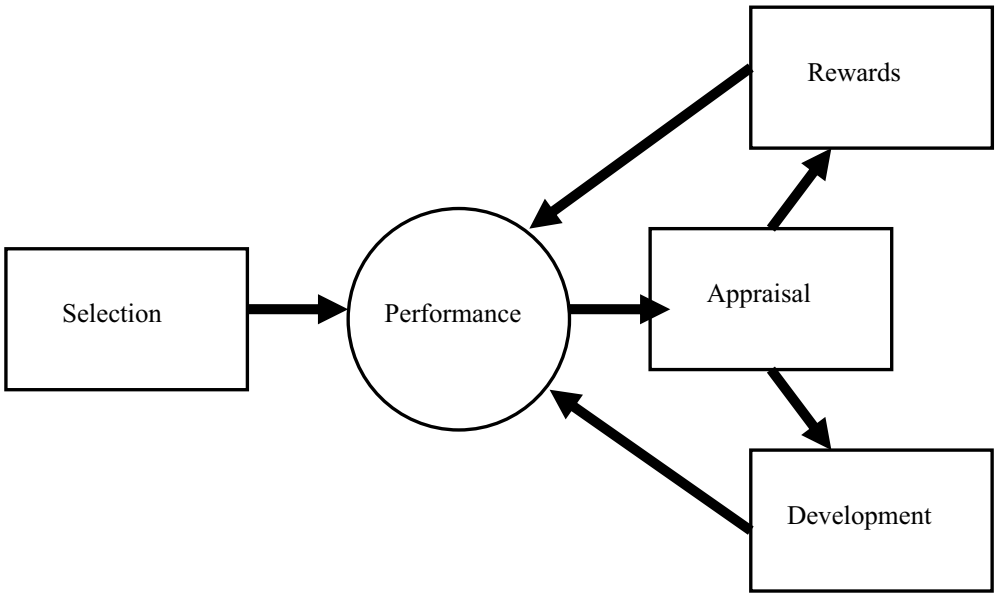

Figure 2. The human resource management cycle

Nowadays, recruitment and selection of staff are needed not only for substituting the employees that have left the organization, or because of upsizing of the organization, they are also seen as an aid and prerequisite for the process of organizational change. Competitive advance is increasingly seen as being based on exploiting and developing the 'core competences' of an organization. The recruitment and selection of managers and other key staff of the organization are of particular importance. It is the competence of managers that will influence the return that an organization will secure from its investments, both in human and material capital (Iles & Salaman 1995). An indispensable part of this managerial competence is to be able to follow the results of essential HRM measures, for example, workload, and to be able to react adequately to these results. According to the idea of HRM, the characteristics of human resources are that they have individual interests, needs and skills and, therefore, they cannot be treated simply as interchangeable numbers, only as a quantitative mismatch between supply and demand (Rothwell 1995). It is not only the qualifications, skill and competence of employees that are important. Personality also has its

own significance within the job and work community (Viitala 2006b). Therefore, in HRM, greater attention is directed at personality profiling.

In recruiting and selecting managers within the health care sector, it is not enough for a person to merely have good professional competence. It is vitally important that candidates also have a sufficient knowledge basis and competence concerning the administration of health care. No-one can be a good manager without this knowledge basis (Kinnunen 2007). Promoting learning and the training of personnel is regarded as central to anything that can sensibly be termed HRM (Keep 1989). It can also be considered as a strategic choice whether an organization increases and develops its core competence mainly through the process of recruitment or training. The principles of HRM also necessitate that politics and practices are designed to realize the latent potential of the workforce at all levels.

The critical factor that distinguishes a thriving organization from one that destroys is its ability to learn from experience and to change, to absorb new information. Since the 1970s, these concepts of the learning organization and organizational learning have become of central interest, stimulated by Chris Argyris and Donald Schön's work (Sädevirta 2004). In organizational learning, organizations observe their own behaviours and actions, and modify them based on the feedback they receive from the environment with the natural goal of improving their performance. Learning consists of two kinds of activity. It can be obtaining know-how to solve specific problems or establishing new premises: paradigms, mental models or perspectives (Argyris & Schön 1978). Peter Senge (1990) recognized that many organizations suffer from 'learning disabilities'. According to him, an organization's commitment to a capacity for learning can be no greater than that of its members. Sydänmaanlakka (2003, 2004) presents the concept of an intelligent organization. It is efficient, capable of learning and sensitive to the well-being of its personnel all at the same time. Thus, it could be characterized as a learning organization in the context of strategic HRM.

Knowledge management (KM) is a multidisciplinary concept with many definitions and many translations, for example, into Finnish (Viitala 2006a). Ikujiro Nonaka is one of the most famous researchers of the knowledge management concept. With their famous book 'The Knowledge Creating Company' (1995) and their other publications, Nonaka and Hirotaka Takeuchi have deeply influenced the theory of knowledge management. They separate two

types of knowledge, explicit and tacit knowledge. Explicit knowledge can be articulated in formal language (grammatical statements, mathematical expressions, and so on) and, thus, can be easily and formally transmitted across individuals. Tacit knowledge, on the contrary, is hard to articulate with formal language. It is personal knowledge embedded in individual experience and involves intangible factors such as personal belief, perspective and value system. The interaction of these two types of knowledge, called knowledge conversion, gives rise to the following four modes: socialization, externalization, combination and internalization. Only if the knowledge is shared with others, is an amplification of knowledge at the group or division level possible, and the knowledge spirals itself organizationally. Although the core of the organizational knowledge-creating process takes place at the group level, the organization provides the necessary enabling conditions. This process is nonlinear and iterative and requires suitable organizational conditions. The most suitable management process for this is so-called middle up-down management. The optimal organization structure for this is called the hypertext organization, a combination of formal hierarchy and flexible task force.

The organizations in the health care sector can be considered as knowledge-intensive service organizations, in which experts perform knowledge work (Pyöriä *et al*. 2005). Intellectual capital is a concept that is quite close to knowledge management. It can be defined as the sum of everything everybody in a company knows that gives that company a competitive edge in the market place (Stewart 1997, Sveiby 1997, Ståhle & Grönroos 2000). According to Stewart, it can be further divided into human capital (employees and their knowledge; such as nurses at a hospital), structural capital (everything that is left at a workplace after a workday; such as PCSs at a hospital) and client capital (customer connections). These three concepts, knowledge work, intellectual and human capital emphasize the central significance of human resources and knowledge of these for the success of an organization. They also help to understand that measuring nurses' workload cannot be simply reduced to measuring their mechanic work performances, but more intellectual performances, for example, monitoring and puzzling over the state of a patient and planning what to do should also be included in the measurement tool of nurses' workload. Furthermore, the amount of nursing workload needed to be allocated to a certain health problem should be considered as a dynamic factor. The concepts of organizational learning and knowledge management also intrinsically include the development of the processes of patient care. Thus, alongside the development of the whole

processes, nursing care practices should also change – changing in turn the workload needed to respond to various health problems.

3.2.2 The roles of management, work organization and measurement in HRM

In HRM, the role of managers is considered critical; they must actively participate in the human resource practice of the organization, along with personnel specialists. As a rule, much greater attention is paid to the management of managers themselves (Storey 1995). According to the principles of HRM, the organization of labour is based on teams. Empowerment means greater responsibility also for those at lower levels (Rothwell 1995).

An old business maxim suggests, "You cannot manage what you cannot measure". The measurement of human resources is complex, difficult and, at times, confusing, but it can and must be done. The process of finding appropriate HRM measures starts from a clear understanding of the organization's goals. These must be converted to measurable HRM practices. Such efforts focus attention on what HRM practices, professionals and departments must deliver to the strategy and for the success of organization. These answers must then be conceptualized, defined – and operationalized for measurement needs (Ulrich 1997). A well-validated PCS, such as the RAFAELA system, for example, fulfils the criteria of a useful measure for the HRM purposes of nurses.

In the health care sector, the role of managers has become increasingly challenging in recent years. Commonly, problems concerning management in health care are proposed as a central cause of many problems in the organizations, although absence of or deficiencies in formal administrative competence are decreasing among health care leaders. It is apparent that problems concerning the management of hospitals and of primary health care are increasingly due to problems in the management system and structure, and not always to a lack of competence or other properties of individual leaders (Kinnunen 2007). The challenges and demands of management competence are increasing and becoming more complex in health care, which finds itself in turbulent conditions (Kinnunen & Lindström 2005, Stenvall *et al*. 2007).

3.2.3 The impact of HRM on the performance of the organization

In his review, Dave Ulrich (1997) analyzes the literature of HRM and comes to the conclusion that solid evidence exists that investment in human resource practices impacts on business results, both financial results and the market value of firms. The alignment of human resources and strategy has an impact on business performance. Executive attention to human resource practices had a large impact on business results, especially under environmental conditions of high change. Nine years later, in the review article of Becker and Huselid (2006), the central question was no longer whether or not HRM – or SHRM – is useful, but how to generate and sustain those obvious potential returns.

3.3 Sickness absence and welfare as a challenge for HRM

3.3.1 The welfare and stamina of personnel

A contradiction can be seen in the management of employees in capitalist systems committed to the production, realization and accumulation of surplus value (Legge 1989). In the long run, however, exploitation of the workforce is not a successful and productive managerial strategy in Western democracies. Job satisfaction and a good organizational climate, motivated and highly educated workers are important elements in achieving a competitive advance. Thus, contributing to the satisfaction and welfare of employees is an essential element of HRM. Selecting the right employees, making individual work contracts, creating working conditions that are flexible and in harmony with worker's individual needs and his/her working capacity are central in HRM. Assignments must also be redesigned to allow devolved responsibility and empowerment (Storey 1995). Experience of justice and perception of fairness in organizations are an essential part of HRM. Distributive and procedural justice are important determinants of employee behaviours or attitudes at the workplace, and in particular, with regard to the commitment of employees towards organizations (Chang 2002). Advanced and validated PCSs may have an essential role in achieving these goals in nurses' HRM, see Table 11 on page 100.

3.3.2 The sickness absence of personnel

HRM stresses the importance of employees as a resource for achieving organizational goals. Therefore, this important resource must be taken good care of in order to promote job welfare and to prevent sick leave. Absence from work due to sickness, that is sickness absence, is an important occupational problem and has an essential impact on loss of productivity and on costs of health insurance. It has been found that, both for men and women, records of medically certified sickness absence are a more powerful predictor of all-cause mortality than established self-reported health measures and available objective measures of specific physical illnesses and medical conditions (Kivimäki *et al*. 2003). Medically certified sickness absences are also a strong predictor of specific causes of death, such as cardiovascular disease, cancer, alcohol-related diseases and suicide (Vahtera *et al*. 2004), and a risk marker for future disability retirement (Kivimäki *et al*. 2004). Thus, medically certified sickness absence may represent a measure of health and the number of them, for example, per ten person-years, can thus be considered an important and well-validated measurement for HRM purposes. Obviously, some of the periods of sick leave represent voluntary absenteeism not related to physical or mental illness (Vahtera *et al*. 2001), and some employees work while ill and record no absences (Kivimäki *et al*. 2005).

Several physical, psychological and socioeconomic factors have an influence on the employee's decision on whether or not to be absent from work. Sickness absenteeism depends on social factors, such as social relations (Melchior *et al*. 2003), and demographic factors, such as age, gender, educational level and employment grade. Women, for example, have higher rates of spells of sickness absence compared with men. Lower occupational status tends to be associated with higher rates of sickness absence (North *et al*. 1996). Societal conditions, such as social insurance and security, influence sickness absence. In many studies, the psychosocial work environment is shown to predict rates of sickness absence. In the Whitehall II study of British civil servants, low levels of work demands, control and support were associated with higher rates of short and long spells of absence (North *et al*. 1996). Analyses of the French Gazel cohort produced somewhat similar results (Niedhammer *et al*. 1998). In many ways, the principles of HRM support the creation of a favourable psychosocial work environment, as was discussed earlier in this chapter.

Only a few studies have examined the association between sickness absence and employees' workload. Sick-listing among women in the public sector was found to be associated with physical or mental demands at work that were experienced to be higher than one's own capacity (Vingård *et al*. 2005). In a questionnaire-based study of workers in the Sweden Post (Voss *et al*. 2004), an association with sick leave was observed for complaints attributed to heavy, arduous work and, for women, a high total workload. Some physical characteristics of work, such as heavy lifting, are also known to be associated with a higher incidence of sick leave (Voss *et al*. 2001). However, there are no contemporary studies examining the relationship between the level of patient-associated workload and the rates of medically and self-certified sick leave.

3.4 Summary of patient classification systems as a part of HRM in health care

The most fundamental idea in HRM is that, among all the factors of production, it is the human resource, its capacity and commitment, which is of crucial significance for the success of an organization. The concept of strategic human resource management (SHRM) can be used if the integration of the human resource function with strategic decision making is especially emphasized. In HRM thinking, personnel are not primarily considered a variable cost to be minimized, but rather a resource for achieving the organization's aims.

The most essential features of HRM can be listed as follows: 1) recruitment is important; 2) performance, not hierarchical position, determines the reward; 3) individual work contracts; 4) division of labour is based on teams; 5) solutions to conflicts between individuals are sought from the organizational climate and culture; 6) working conditions are in harmony with the worker's individual needs; 7) promoting learning and training is central, 8) flexibility is emphasized in following rules and 9) measuring is important.

The main parts or interventions of HRM are the four key elements of the HRM cycle: selection, appraisal, development and rewards. An organization must recruit and select – and train/develop – its employees very carefully to be able to respond to its changing demands of core competence corresponding with its strategy. Organizational learning is an important element in their high performance.

The measurement of human resources is often complex and difficult, but it must still be done to enable successful management. The measures must support the organization to achieve its central goals. An optimal workload can, for example, be considered one such central goal in the HRM of nurses and other employees. The organizations in the health care sector are knowledge-intensive service organizations, thus, knowledge management and intellectual capital are central concepts.

Job satisfaction and a good organizational climate, motivated and highly educated workers are important elements in achieving a competitive advance. Medically certified sickness absence may represent a measure of health and the number of them can thus be considered an important and well-validated measurement for HRM purposes.

4 Previous studies on patient classification systems, workload, staffing and validation of measures in health care

4.1 Testing the validity of measures in health care

Testing the validity of a scale means determining whether the scale measures what it was intended to measure. In natural sciences, the problems of measurements' validity are usually less insignificant than in social sciences, in which measurement results of various factors are more dependent upon how they are defined and how they are measured. In addition, the relation between observation and what it reflects is not always as unambiguous in social sciences (Streiner & Norman 1995, Metsämuuronen 2003).

Validity is not solely a feature of a test or an instrument. It is also always related to the population concerning which the instrument is employed. Reliability also places an upper limit on validity. The types of validity most frequently mentioned are content, criterion-related and construct validity. However, rather than being disparate attributes, the various 'types' of validity all assess the same issue of the degree of confidence that can be placed in the inferences drawn from scores or scales. (Giovannetti 1979, Streiner & Norman 1995).

Content validity indicates how comprehensively and representatively the selected indicators and the items they comprise describe the phenomenon studied (Reitz 1985, Gallagher 1987, Hoffman 1988, Williams 1988, Streiner & Norman 1995). Content validity is frequently studied in the form of expert estimates of clarity and completeness. In order for the indicator to be valid from the content point of view, it should apply to all the relevant parts of the area measured. For its part, content validity reveals the sensitivity of the measuring instrument and comprises a subjective estimate of the measurement rather than statistical analysis (McDowell & Newell 1996).

Within criterion validity, Streiner and Norman (1995) distinguish concurrent and predictive validity. When studying criterion-related validity, statistical methods are most often used, and attention is focused on the validity of the measuring itself. An acceptable procedure is also the

parallel use of a 'gold standard', with which the new method of measuring is compared (McDowell & Newell 1996). The problems faced with regard to the testing of criterion validity are (a) if a better measure, a 'gold standard', already exists why make a worse measure and (b) if the new measure were better, if the relationship between the old and new measure is less than perfect, which one is at fault (Streiner & Norman 1995)? Concurrent validity can be assessed by testing an old instrument against a new one in order to establish whether the new instrument has diagnostic utility and can thus replace the already tested one. An important and sometimes difficult question is whether the instrument used only for the comparison will prove valid in a new organization as well (Giovannetti 1979, Ebener 1985).

Predictive validity indicates the ability of the instrument to predict. As far as instruments for patient classification are concerned, the ability to predict staff resources is in focus (Gallagher 1987). Observational studies concerning the nursing care that patients have received will then concern whether the amount of care that has been established for the category in question has actually been provided (Giovannetti 1979, Williams 1988). Another method consists of validating by means of time studies, that is, by the ability of the instrument to predict how much of the nurses' time patients in different categories will need (Giovannetti & Mayer 1984, Williams 1988, De Groot 1989). Predictive validity testing is seldom used, because the method is considered to involve logistical problems (McDowell & Newell 1996).

Construct validity has received little attention (Alward 1983, Ebener 1985, Whitney & Killien 1987) in research on patient classification. It can be used even in such situations when no other similar type of measure exists to compare it with. Construct validity concerns the ability of the instrument to measure hypothetical constructs; in other words, the variables used in an instrument can be seen as theoretical abstractions (Whitney & Killien 1988). These hypothetical constructs can be seen as a mini-theory which explains the connections between the variables in the phenomenon studied, i.e., the variables are conceptually linked to form 'syndromes' (Streiner & Norman 1995). Construct validity describes how well the instrument is operationalized and quantified; in other words, it testifies that coefficients and other quantitative aspects are adequate (Ebener 1985). There is no one single experiment that can unequivocally 'prove' a construct. Thus, construct validation is an on-going process. For assessing construct validity, two general strategies are used, that is, by searching for the convergent validity and the discriminant validity (Woods & Catanzaro 1988). Discriminant and convergent validity are usually assessed using a powerful technique, the so-called

multitrait-multimethod matrix (Streiner & Norman 1995); this technique, among others, was used for example in Hinshaw and Atwood's (1982) study.

De Groot (1989) sees the choice of critical indicators as decisive with regard to validity but, at the same time, emphasizes that a greater number of indicators has not been proven to increase validity. Conversely, conceptual relevance and the simplicity of the instrument are important. The question of validity, whether instruments in fact measure what they are designed for (Ebener 1985, Hoffman 1988), is of great importance with regard to their reliability. However, because the conception of caring has undergone changes in nursing science (Strandmark Kjölsrud 1995, Tuomi 1997), the importance of accurate and continuous evaluation of validity must still be stressed. Since validity is not exclusively a feature of a test, but also depends on the population, it must always be reassured and retested when instruments of measurement, such as patient classification systems, are used in a new environment.

4.2 Testing the reliability of measures in health care

Reliability can be considered an index of the extent to which measurements of individuals obtained under different circumstances yield similar results. Because of the limited prior information on the true variation between individuals, reliability is usually quoted as a ratio of the variability between individuals to the total variability (the sum of 'true' score and 'error' score) in the measurements. Thus, reliability can be said to be a measure of the proportion of the variability in scores, which is due to true differences between individuals, and it is expressed as a number between 0 and 1 so that zero indicates no reliability, one indicates no measurement error and perfect reliability (Streiner & Norman 1995).

There are several different ways to measure reliability. (a) Internal consistency measures if the different items measure the same thing; it represents the average of the correlations among all items in the measure. Cronbach's alpha, for example, is such a measure. However, internal consistency as such, is not a sufficient basis upon which to make a reasoned judgement. (b) Stability examines the reproducibility of a measure administered on different occasions. The different methods are inter-observer, intra-observer and test-retest reliability. There has been considerable debate in the literature regarding the most appropriate choice of reliability

coefficient. Percentage of agreement is widely used in standard practice with PCSs. Cohen's (1960 in Streiner & Norman 1995) weighted kappa coefficient, for example, is more exact than percentage of agreement (Streiner & Norman 1995, Metsämuuronen 2003).

Reliability is always related to the population. There is literally no such thing as the reliability of a test. Moreover, according to the generalizability theory, the classical conceptualization of measure as a sum of 'true' score and 'error' score can be considered as an oversimplified assumption (Streiner & Norman 1995).

4.3 Studies on workload, staffing and patient classification systems

The different types of measurement systems for assessing staffing adequacy and determining nurses' workload are presented in Chapter 2.3. Measurement systems for nurse requirements and nurses' workload or nursing intensity should be tested for reliability and validity – and feasibility – before being promoted or adopted on a wide scale. Likewise, different measurement systems should be compared with each other in order to discover if they produce comparable results. In principle, the quality of a measure depends both on the quality of the input data and on the performance of the measurement instrument itself.

The reliability of any systematic recording of nursing activities and workload ultimately rests with the ward nursing staff. Failure to train, educate and explain the necessity for resource management results in failures to accept the importance of data collection. Perceptions of the misuse and manipulation ('acuity creep') of systems to inflate staffing needs are said to be not uncommon (Van Slyk 1991, Malloch & Conovaloff 1999). Costly and increasingly adversarial audit processes have been incorporated into many organizations to monitor and improve patient classification compliance among nurses to better control resource use (Shaha 1995). A much greater problem than possible manipulations is a careless attitude towards documenting; 'we do the entry when we have time'. The accuracy of data input may, thus, suffer. The quality of the output cannot be better than that of the input ("garbage in, garbage out").

Some studies have tried to compare different methods of determining nursing workload. In the USA, Schroeder *et al*. (1984) compared different methods that were available at the

beginning of the 1980's and found little difference between the number of nurses determined by category-oriented (bottom-up dependency) and task-oriented (bottom-up intervention) nurse demand methods. Rees *et al*. (1991) compared four workload systems using a quasi-experimental approach. Nurses were asked to assess model patients on four imaginary wards with each ward using one of the different systems. There was a low intra- and inter-system reliability. However, the nurses were not familiar with the four systems.

In the United Kingdom, Carr-Hill and Jenkins-Clarke (1995) compared four nursing workload management systems (C for C, FIP, EXCELCARE, SENS) that represented methods 1, 2 and 3 according to their classification (see Table 2). Data were collected in three wards over six days. The absolute value of the percentual difference of actually measured and predicted working hours was 17% (range 2 – 38%; all figures are calculated from their study report). Seago (2002) compared summative and criterion types of patient classification system, both of which were widely used in California. The summative task type PCS requires the nurse to check off activities, treatments and procedures according to the frequency of occurrence for each patient in her care. The critical incident or criterion type PCS uses broad indicators to categorize patient care activities. There were virtually no differences in the predictive ability of these two instruments in the results. They both predicted exact actual scores about 79% of the time.

In recent years, researchers have studied whether triangulation of the different workforce planning methods would improve the outcome. Ellis and Chapman (2006) describe the 'GOSHman project', which aims to identify nurse staffing requirements based on paediatric patient acuity and dependency. Correlation between predicted and current staffing appeared to be close for most wards and departments, but the need for higher nurse establishments could be identified more clearly and accurately where patient acuity had increased or overall nurse staffing levels were inadequate. Harrison (2004) describes a similar idea for the classification of critical care patients, in which the SUHT acuity and dependency tool measures both acuity and dependency.

It can be concluded from these studies and from some minor studies described in overview articles (Arthur & James 1994, Hughes 1999) that there are only few studies that compare different patient classification systems and, moreover, there are many limitations even with these studies. The results of older studies, in particular, are quite variable and, for the most

part, not especially satisfying. However, the results of the most recent studies are more promising. Using triangulation of the different workforce planning methods may improve the result, but at the expense of increased time and effort.

Nurses' perceptions of staffing adequacy, however, also include some subjective elements. In Mark's (2002) study, it was shown that there is a clear negative association between perceptions of staffing adequacy and the number of beds on the unit. Likewise, higher levels of patient technology were also associated with perceptions of less adequate staffing. According to Hughes (1999), there are several problems in assessing nursing workload. In addition to concerns dealing mainly with nursing philosophy, she states that, in activity-based methods, nursing is seen as a linear activity instead of 'multitasking'. The time can be governed by the patient, not the nurse. Other aspects of nursing behaviour than activities (thinking, skills) can be missed. Tasks are also prioritized and performed in a variety of ways and at different speeds depending on the circumstances.

Procter (1992) criticizes the current tendency of focusing on producing objective measurements of nursing workload. She claims that, in practice, nurses tend to identify only those patient care needs that can be met by the current staffing levels. The categories and timings given to the nurses are derived from observations of their current practice and, therefore, the prevailing staffing structure will tend to dominate any attempts to measure workload. She proposes that identifying appropriate staffing levels on a ward requires careful and painful negotiation which will inevitably incorporate a large element of subjective professional judgement.

In conclusion, even most of those authors who have criticised the available measurement instruments admit that methods for determining nurse staffing requirements are needed to make nursing workload visible. According to Arthur and James (1994), since a perfect workload measurement system is unlikely ever to exist, such systems can be used to facilitate, but not to dictate, decisions about nurse staffing. Hughes (1999) also proposes a pragmatic solution of using simple methods of workload assessment that can be used in monitoring the relationship between skill mix and the quality of nursing care. Needham (1997) concludes that despite the problems of measuring nursing workload, realistically, however, a compromise that uses patient dependency, together with the impartial allocation of resources, is the nearest to an accurate measurement of nursing workload that can be expected. Thus, measurement

systems for nurse requirements and nursing intensity and workload are far from being ideal, but despite their deficiencies they are useful and needed.

4.4 Validation of measurement tools in selected Finnish doctoral dissertations in the field of health sciences

4.4.1 Introduction

The previous sections in this fourth chapter have been devoted to providing a discussion of the problems of testing the validity and reliability of measures in health care at a more general level; for example, presenting the types of validity and reliability and the alternatives in measuring these. In the sections that follow, the focus is directed away from the general level towards specific cases. The object is to study how these questions concerning the validity and reliability of new measurement tools have, in practice, been resolved in Finnish health care. For this purpose, four relevant Finnish doctoral dissertations that had a measure validation in the field of health sciences as their orientation or part of their task were chosen for closer analysis. Thus, the analysis concentrates wholly on methodological items, passing by the scientific results of these publications and their significance concerning health sciences or health care.

For most needs of measurement, it is possible to employ an already existing measure. Sometimes, however, there is no instrument or scale which may suite the purpose. In these situations, a new measure must be created, most often based on one or more existing measures. A more common problem, for example, in Finland, is that an existing measure must be translated and cross-culturally adapted to another language and context.

4.4.2 The background and object of the selected doctoral dissertations

Marja-Leena Perälä's thesis (1995) aimed at testing the Qualpacs quality measurement instrument. This measure was developed on the basis of patients' needs. It consists of 68 items, grouped in six scales. The data of this measure was collected through observation of the nursing care of individual patients in all kinds of situations in which patients and nurses interacted. In this thesis, the measure was first translated into Finnish. Perälä's material

consisted of an evaluation performed by nursing experts and by a sample of a population and by use of the measure in practice.

In her doctoral dissertation (1999), **Ann-Marie Turtiainen** analyzed the validity, reliability and credibility of the Belgium Nursing Minimum Data Set (BeNMDS), which had first to be adapted to Finnish conditions. The nursing minimum data set comprises the most essential data of nursing care and it is collected directly from practice four times a year. This collected data is either in the form of nursing nomenclature or standardized nursing classification. The BeNMDS includes data regarding nursing activities, and also data concerning organization, personnel resources, medicine and activities of daily living (through the San Jose patient classification). The nursing data consists of 23 nursing activities. The object of her research was to study these nursing activities in this BeNMDS-system.

Päivi Voutilainen's thesis (2004) intended to provide information on the quality of nursing care in the long-term institutional care of older people by examining the structure, process and outcome dimensions of quality. The quality dimension of the nursing process was assessed by studying nursing documentation. The required nursing data were collected from patient records. This data set was gathered using the Senior Monitor quality assessment instrument. Before this, the psychometric properties of this instrument were determined. First, however, the instrument was translated from English to Finnish.

Juha Laine's thesis (2005) studied the connection between productive efficiency and care quality in institutional care for older people in Finland. His thesis belongs to the field of health economics and productivity research. His material consisted of 11% of all Finnish older people in institutional care in the wards of residential homes and municipal health centres. The RAI system was used as a data system. It is a comprehensive international system for assessing and following the care quality of the elderly. It comprises the whole of the client's care and service chain and is integrated as part of a patient administration system. It comprises clients' minimum data set, a manual and measuring instruments (client mix, quality indicators, etc.). The validity and feasibility of the RAI system has been widely studied in many countries, with good results. Thus, there was no need for further testing of these aspects. Productive efficiency was expressed as a ratio of the observed productivity of an organization to the best possible productivity of the whole material. The case mix of the clients was standardized by RUG-III/22-classification. The quality of care was

operationalized to clinical quality indicators that were received from the RAI database. For practical reasons, the concept of quality was operationalized mainly to the occurrence of detrimental and unwanted processes and outcomes of care. The client's point of view could not be used because most of the clients suffered from cognitive decline.

4.4.3 Testing and estimating validity in the selected doctoral dissertations

In Marja-Leena Perälä's thesis (Table 5), evaluation of the content and the construct validity was performed by nursing experts. A random sample from the population was also used for testing the content validity. Both nurses and people in the population sample evaluated the importance of different items as part of the content validity of the measure. Nurse experts also evaluated whether various items could be considered as a quality indicator of patient care. Testing of the content and construct validity and reliability was also performed through observations of patient care situations. The relevance of different items as part of the content validity was evaluated. Factor analysis was used in defining the conceptual structure of the measure.

In Ann-Marie Turtiainen's thesis, the content validity of the BeNMDS was tested using content analysis of Finnish nursing doctoral dissertations and Katie Eriksson's publications of her Nursing Process Model. In addition, focus-group interviews of the health care administrators were also used to test further the content validity. The testing of construct validity was done by testing factorial validity. All the validity, reliability and sensitivity tests mentioned here were performed using the data collected with the BeNMDS-tool from patients' nursing notes in Finnish Hospitals.

In Päivi Voutilainen's thesis, the psychometric properties of the Senior Monitor quality assessment instrument were determined. The content validity was assessed by an expert panel.

In Juha Laine's thesis, the validity and feasibility of the RAI system were considered. The central issues of validity concerned: (1) how valid the measurement of productivity of older people's institutional care at various institutions was, and whether there were other ward-level variables that could be approved; (2) the operationalization of the concept of quality to the occurrence of detrimental and unwanted processes and outcomes of care. It would have been

favourable to use a wider concept of output quality, including variables of effectiveness, but for practical reasons, this was not possible.

Table 5. The validity and reliability testing of selected Finnish doctoral dissertations in the field of health sciences with a measure validation as their orientation or part of their task.

Doctoral dissertation	**Analysis of validity and statistical methods in it**	**Analysis of reliability and statistical methods in it**	**Other analyses**
Perälä, Marja-Leena. "The quality assessment of patient care: validation of a quality measure (Qualpacs)."	Content validity -nurse experts -population sample -observations of patient care Criterion validity -no analysis Construct validity -nurse experts -observations of patient care (factor analysis) -population sample (factor analysis)	Stability -no analysis Equivalence -Percentage agreement -Cohen's kappa -intraclass correlations Internal consistency -Cronbach's alpha	
Turtiainen, Ann-Marie: "Methods to describe nursing with uniform language: The cross-cultural adaptation process of the Belgium Nursing Minimum Data Set in Finland."	Content validity -content analysis of nursing dissertations -content analysis of Katie Eriksson's publications -focus-group interviews of health care managers Criterion validity -no analysis Construct validity -factorial validity using nursing data	Stability -test-retest (ridit, OR) Equivalence -inter-rater reliability (lambda, Kendall's tau, Pearson's r) Internal consistency -principal component analysis	Sensitivity -visual testing
Voutilainen Päivi. "The quality of nursing care in the long-term care of older people. (Analyses here are related to the Senior Monitor quality assessment instrument)"	Content validity -expert panel Criterion validity -no analysis Construct validity -no analysis	Stability -no analysis Equivalence -inter-rater reliability (Percentage agreement) Internal consistency -Cronbach's alpha	
Laine, Juha. "Quality and Productive Efficiency? An Economic Study on Institutional Care for Older People."	See text	See text	

In conclusion (Table 5), content validity has been tested frequently, using expert panels or focus group interviews. Population sample was used in Perälä's thesis, but patients'/clients' viewpoints were not utilized, as was the case, for example, in the validation of the RAFAELA (Fagerström *et al*. 1999). This can be said to have been impossible in the case of Voutilainen's and Laine's dissertations, because many of the clients suffered from a cognitive dysfunction. Nor has any use been made of criterion validity at all, because of the lack of an instrument that could have been considered as a 'gold standard'. Various methods belonging to factor analysis have been used in the analysis of construct validity.

4.4.4 Testing reliability in the selected doctoral dissertations

In Marja-Leena Perälä's thesis, percentage agreement, Cohen's kappa and intraclass correlations were used in testing the equivalence of measure, and Cronbach's alpha in measuring the internal consistency of measure. In Ann-Marie Turtiainen's thesis, testing the reliability of the tool was done by separate testing of its stability, equivalence and consistency. The stability was tested using the test-retest technique, and the equivalence using inter-rater techniques. The consistency, the construct validity and the weight of the dimensions of the tool were tested using principal component analysis. The 'fingerprints' of the wards and the 'national nursing map' produced by the BeNMDS-program were used to test the sensitivity of the tool in Finnish nursing practice. In Päivi Voutilainen's thesis, the equivalence of the Senior Monitor quality assessment instrument was assessed using inter-rater percentage agreement. The internal consistency was tested using Cronbach's alpha.

In conclusion, testing of equivalence was employed more often than testing of stability, for known reasons (Turtiainen 1999, Metsämuuronen 2003). The phenomenon and objects that were measured, for example, should have remained unchanged between repeated measurements: a presumption that is most often incorrect. The internal consistency of the measure has been commonly determined using different methods.

5 The RAFAELA system

5.1 General presentation of the RAFAELA system

Patient classification systems and their creation were presented in Chapter 2 in general terms. Likewise, the validity and reliability testing of PCSs and other measures in health care were presented in the previous chapter in general terms. In this chapter, all these items are discussed in relation to a case, the RAFAELA system: the patient classification system that is by far most widely and still increasingly used in Finland. The chapter begins with a general outline of this rather complicated PCS in the form of an illustrative example, followed by a more in-depth presentation of its components and details. Its development process is then discussed. The chapter closes with a description of the validation and reliability testing processes concerned with the system.

5.1.1 The parts of the RAFAELA system

1. All patients' nursing intensity is measured daily using the **OPC**.
2. Data of the **daily nursing resources** is collected.
3. The optimal nursing intensity as OPC points per nurse is periodically assessed using the **PAONCIL** as a comparison.

5.1.2 The basic idea of the RAFAELA system in the light of an example

Phase 1: (See Figure 3.) **Every patient's nursing intensity** is measured daily using the OPC. A patient can receive 1 – 4 points from each of the six subsections, thus in total 6 – 24 OPC points. In the example ward, there are only seven patients, and the daily OPC nursing intensity of each patient is determined. In this particular ward on that day, it varies between 12 and 21 OPC points per patient. **The total nursing intensity of the ward** is then calculated by summing up the total OPC points of all individual patients and, in this example, it is 100 OPC points.

Phase 2: **The nurse resources** are determined. Nurse resources mean the total number of nurses who have nursed the patients in the ward during that calendar day; in this case 5 nurses.

Phase 3: The daily **nursing intensity per nurse** is then calculated by dividing the whole ward's total nursing intensity (OPC point sum) by the number of nurse resources on that calendar day. In this case the result becomes 20 OPC points per nurse.

Phase 4: Periodically, for example, every other year, the **PAONCIL** instrument is used simultaneously with the other two measures for some weeks **in order to estimate the nurses' daily workload in relation to an optimum** in that ward. During that period, each nurse assesses every shift/daily on a scale of -3 to +3, as to whether they considered their patient-associated workload as optimal (= 0), above optimal (>0) or below optimal (<0).

Phase 5: The OPC/nurse values are then compared to the average PAONCIL values of the same day using ward data from a period covering a few weeks. In this way, using **linear regression analysis**, the optimal nursing intensity per nurse value of that ward is then finally determined, that is, the OPC/nurse value that corresponds to the average PAONCIL value zero of the same day. In this example, the ward's optimal nursing intensity was 18 OPC points per nurse. The optimal range (or level) is in practice defined as optimal point ±15 %.

Phase 6: When the ward's optimal OPC points per nurse value and its range are known, the actual nursing intensity per nurse (on the day of the example 20 OPC points per nurse) can be compared to the optimal points and range every day.

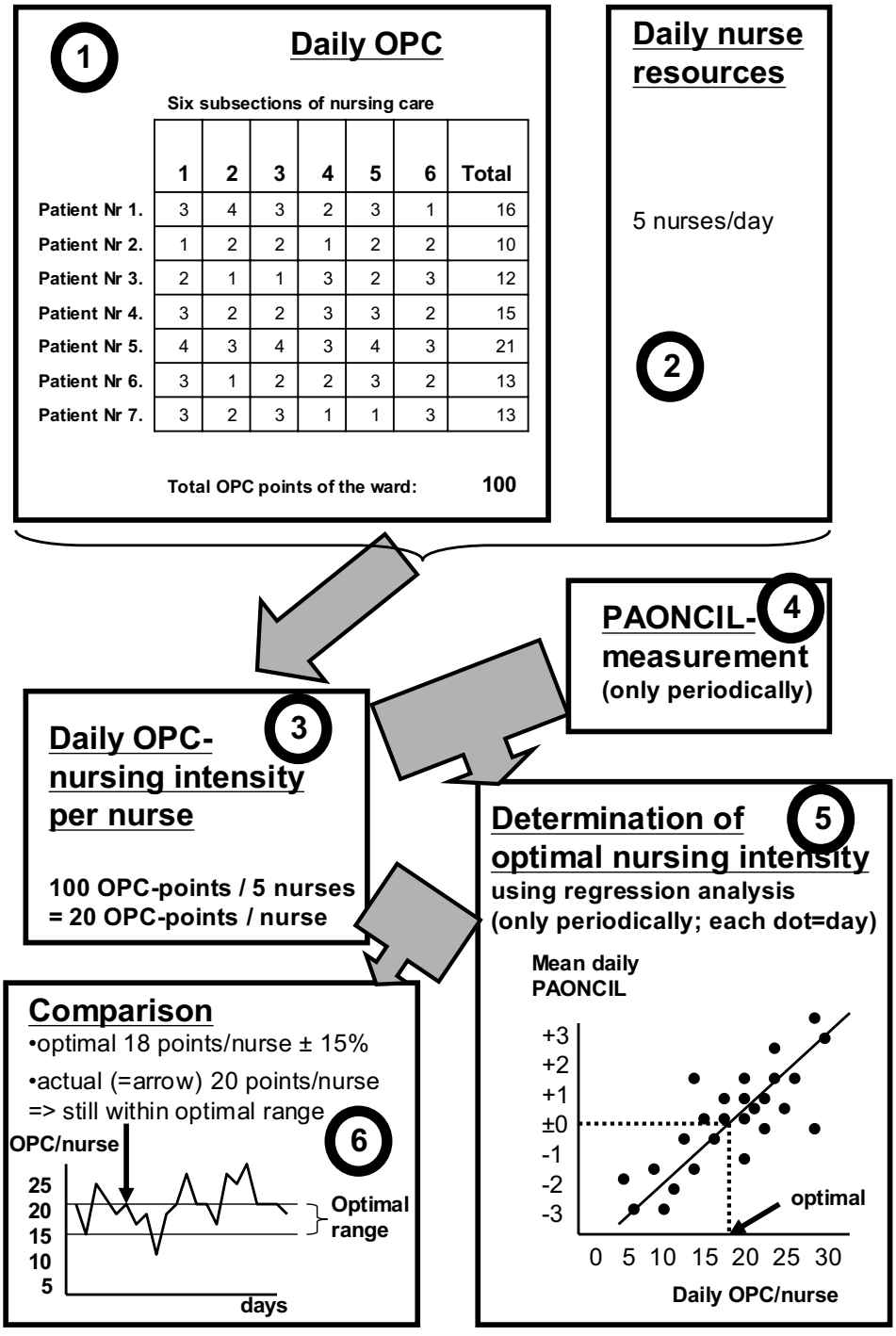

Figure 3. Presentation of the main parts of the RAFAELA system in the light of an example.

The number of patients and nurses is small for the sake of simplicity. For explanations of the various phases 1 – 6, see Chapter 5.1.2.

5.2 Closer presentation of the RAFAELA system

According to the **OPC** (Figure 4.), nursing care and the caring needs connected with this consist of the following six subsections: 1. Planning and co-ordination of nursing care; 2. Breathing, blood circulation and symptoms of disease; 3. Nutrition and medication; 4. Personal hygiene and secretion; 5. Activity, sleep and rest; and 6. Teaching, guidance in care and follow-up care, emotional support (Rainio 1994, Kaustinen 1995, Fagerström *et al*. 1998, Fagerström 1999). These six subsections of nursing care contain different nursing care activities through which patients' caring needs are met.

The six subsections of nursing care in OPC

Nursing intensity	1.	2.	3.	4.	5.	6.
A	1	1	1	1	1	1
B	2	2	2	2	2	2
C	3	3	3	3	3	3
D	4	4	4	4	4	4

Figure 4. Presentation of the main parts of the OPC nursing intensity measurement instrument.

A patient can score 1 – 4 nursing intensity points from each of the six subsections. Thus, total nursing intensity points per day can vary between 6 and 24 points.

The first phase in the RAFAELA system is – in accordance with the written OPC instructions – that every day, the nurse classifies electronically the patients they have nursed. The nurse

makes the classification on the basis of care actually provided, not on that of desirable care (Hoffman 1988). Nursing intensity can vary for each area from A (= 1 point), B (= 2 points) and C (= 3 points) to D (= 4 points). A identifies a patient who manages relatively well on his/her own; B a patient who is occasionally in need of care; C refers to repeated need for care; and D refers to the caring needs of a patient who cannot manage unaided at all. The points are added up, forming 6 – 24 points per patient. The total sum of nursing intensity points for the ward is then calculated.

The next phase in the RAFAELA system is that the total sum of OPC nursing intensity points for a ward, for example, 200 points, is divided by the total number of nurses (for example 10 nurses) who had nursed the patients in the ward during that calendar day. The actual patient-associated workload might then be, for example, 20 nursing intensity points per nurse.

The next phase consists of a **PAONCIL** questionnaire period. Before the studies of this thesis, this period varied from six to eight weeks. The professional assessment of nursing intensity proceeds as follows: after each work shift, each nurse records on a form a numerical estimate (on a scale from minus 3 to 3, with an accuracy of 0.25 points) of the extent to which she/he had time to meet the caring needs of the group of patients she/he had taken care of during the shift. The scale comprises the following levels of nursing intensity: 3 = very high; 2 = high; 1 = fairly high; 0 = optimal level, -1 = fairly low; -2 = low; and -3 = very low (Fagerström & Rainio 1999, Fagerström 1999). Guidelines for the assessment are presented in a manual, in which the seven levels of nursing intensity are described in relation to several important quality indicators of nursing care.

The descriptions of these seven nursing intensity levels also contain quality aspects. Thus, for example, very high nursing intensity implies a situation in which the nurses are unable to respond adequately to patients' caring needs, i.e., the nursing intensity level is too high in relation to staff resources. The result is that the care takes the form of meeting only the most urgent needs of patients. Information to patients and their next of kin remains insufficient and the planning of care defective. The risk of mistakes increases and tasks are postponed until the next shift. The 0-level, representing optimal nursing intensity, is defined as a situation in which the patients receive good holistic nursing care – physical, psychological, social and spiritual. It is then possible for the nurses to fulfil the patients' caring needs and also to guide, support and inform the patient.

The aim of the PAONCIL method is – according to the best practices of HRM – to ascertain the **optimal nursing intensity level** for each ward in OPC points per nurse. According to the basic idea of the RAFAELA system, the daily patient-associated workload in OPC points per nurse should be compared with the optimal nursing intensity level for each ward, which is also defined as a certain level of nursing intensity points per nurse. In practice, it has been decided that the optimal level is optimal nursing intensity point ± 15%.

Optimal nursing intensity is produced by simple linear regression analysis (Figure 5). The linear association between the values of the OPC (daily points per nurse) and the PAONCIL (daily mean) instruments can be quantified as follows: what value does the OPC give when the average PAONCIL for the same day is optimal (i.e. zero), and how strong is the association (the explanatory power) between the OPC and PAONCIL? The explanatory power – or the determination coefficient (R2) – determines by how many per cent the variation in values of the OPC explains the variation in values of the PAONCIL. This can, in principle, vary between 0% and 100% (Fagerström & Rainio 1999, Fagerström 2000, Fagerström *et al*. 2002).

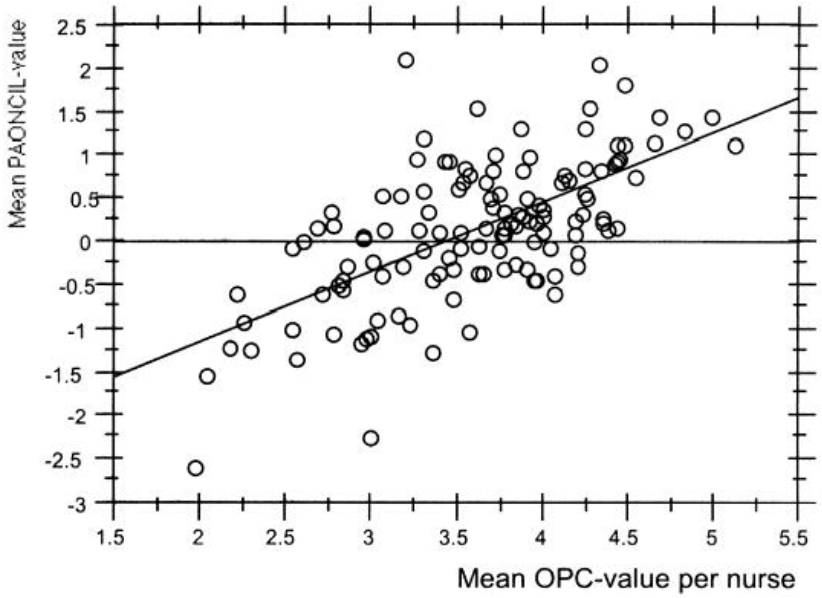

Figure 5. A scattergram of the relationship between the daily mean values of OPC per nurse and the daily mean values of the PAONCIL, with the fitted line and 95% confidence bands of the mean.

The figure presents the data from an internal medicine ward. The corresponding regression equation is y = 0.804 *× 2.761; F = 86.040; P < 0.0001; R^2 = 0.408. Optimal nursing intensity is ca. 3.5 points/nurse. The numerical values of OPC points are unexpectedly low because class coefficients are used instead of OPC points, as was praxis earlier. This figure is from the first sub-study and is published with permission of Scand J Caring Sci and Wiley-Blackwell Publishing.

The RAFAELA system is difficult to place in Malloch and Conovaloff's (1999) classification of patient classification systems in three generations (see Chapter 2.3). It is no doubt an advanced system using the systems of information technology. Conversely, there are, for example, no direct economic and outcome measures and it does not report the skill mix needed. The system can, of course, be integrated with other data systems to produce all such parameters – and this is routinely done in some hospitals in Finland. It can be considered a matter of taste, whether all the attributes required from a third generation PCS for the HRM and financial management of nursing are included in the same monolithic system or if they are produced by separate systems integrated to the PCS.

5.3 The development process of the RAFAELA system

5.3.1 The development process of the OPC

The Oulu Patient Classification (OPC) has been developed on the basis of the Canadian HSSG (Hospital Systems Study Group) classification in Finland, at the Oulu University Hospital. Patient classification according to the HSSG method stressed basic physical needs and the care these presupposed. Nursing care planning and co-ordination, rehabilitation, and information and guidance to patients and their closest relatives were lacking from the HSSG method, which therefore had to be revised to become more suitable for individual and holistic care and more compatible with the prevalent Finnish caring ideology (Kaustinen 1995).

The basis of the OPC are the principles of nursing care presented in the quality assurance programme of the neurology ward in the neurological clinic of Oulu University Hospital in Finland and Roper's model of nursing care. During 1991-1993, a work team participated in the development of the OPC. This work team comprised special nurses, ward nurses and head

nurses from six different wards, representing the special fields of surgery, internal medicine, neurology and paediatrics; the team also included information technology experts. The validity and reliability of the classification were tested (Kaustinen 1995).

At the Vaasa Central Hospital in western Finland, the implementation of a new method of patient classification, the Oulu Patient Classification (OPC), was started in 1995 (Fagerström & Rainio 1998, Fagerström 1999). It was soon noticed that the mere division of patients into four or five nursing intensity categories based on their total daily OPC points did not provide enough information on nurses' workload.

The following step was to calculate the weighting coefficients of the various categories. These were created at the Vaasa Central Hospital. The mean OPC points of each of the categories were calculated, and these mean values were then divided by the mean of the first category, that was, by seven. Then, the first category was given the weighting coefficient one and, with this as a basis for comparison, the coefficients for the other categories were established. By summing up the coefficient points of all patients in the ward, the total nursing intensity score of the ward was then obtained. Later, it was realised that dividing patients into different categories was an unnecessary intermediate step, which only led to an unnecessary loss of some information. The total nursing intensity score of the ward could be counted directly using the sum of the six subsection's OPC points of every patient. In practice, the ratio between direct OPC points and coefficient points is approximately 6.73:1 (Fagerström & Rauhala 2001). The OPC manual was further developed based on the results of expert validation (Fagerström 2000).

However, the total nursing intensity score, which indicated the direct and indirect patient-associated nursing workload of a ward, caused by the patients' caring needs during a calendar day, was not by itself sufficient, because it did not take into account the number of nurses. For this reason, it was decided to divide the total nursing intensity points for a day by the total number of nurses in patient care on the morning, evening and night shifts. The daily patient classifications thus result in a daily nursing intensity level expressed in OPC points per nurse.

5.3.2 The development process of the PAONCIL

The development of nursing intensity levels, expressed in terms of nursing intensity points per nurse, was not, however, enough. Further questions arose. What might the optimal nursing intensity level per nurse be? How high could nursing intensity per nurse be without endangering the quality of the care and risking that the patients did not have the possibility of getting their caring needs met (Fagerström & Rainio 1998, Fagerström 1999)?

A method was called for which did not demand too many resources, could be easily implemented, and periodically be applied daily, and did not require too much additional work from the nursing staff. It was also found that it is important to start from the experts and staff when considering the need for resources for the good care of patients, thus proceeding from the bottom-up idea. The importance of subjective or professional estimates should not be undervalued when the adequacy of staff allocation is tested. The validity of patient classification is called into question if statistical data are in conflict with the assessment of experts and staff (Arthur & James 1994).

The PAONCIL method has been under development at the Vaasa Central Hospital since 1995. It was considered to be an alternative to classical time studies. These were developed under the influence of a technological, industrial view of nursing and they have been considered as reducing nursing to a set of nursing acts. The PAONCIL method is an example of a method which has been developed on the basis of a caring science perspective (Fagerström & Rainio 1998, Fagerström 1999). The PAONCIL method is based on an HRM style of leadership, which presupposes reliance on the competence, sense of responsibility and power of judgement represented by professional nurses when the situation is such that it is realistically possible to give patients good care and make caring the essence of nursing (Fagerström & Rainio 1998, Fagerström 1999). When nurse managers show confidence in the nurses' competence, this has a strengthening and encouraging effect on nursing (Brown 1991). The method is based on the interactive and dynamic character of caring and presupposes that nurses take responsibility for patient groups and, in this connection, prioritize patients' caring needs, something that De Groot (1994a, b) also stresses in the development of 'The Patient Classification System Staffing Matrix'. Using the daily OPC points per nurse and average PAONCIL value, it became possible to determine the optimal

nursing intensity of different wards expressed as OPC points per nurse by means of simple linear regression analysis (Fagerström & Rainio 1999, Fagerström 1999).

5.3.3 The national development process with the RAFAELA system

The RAFAELA system was reported as being a useful tool in personnel administration (Rauhala *et al*. 1999), and it was presented at many educational events. The situation was favourable for a new innovative patient classification system. There was a clear need for a measurement system and, at the end of the 1990s, it began spontaneously to spread to other hospitals.

The spontaneous and non-coordinated spreading of the RAFAELA system led to a need to determine its value and usability, before it became a *de facto* standard uncoordinatedly. Then, under the auspices of the Association of Finnish Regional and Local Authorities and good offices of development manager Ritva Larjomaa, a comprehensive research project was carried out in 2000 – 2002, testing how useful and workable the RAFAELA system was in Finnish health care. The results were considered positive and encouraging (Fagerström & Rauhala 2001, Fagerström & Rauhala 2003). As a consequence of this co-operative project, the RAFAELA system continued to spread, and it became obvious that there must be an organization to administer and coordinate the use, uniformity and development of this system.

To ensure the correct and uniform use and reliability of the OPC and the PAONCIL measures of the RAFAELA system, in 2003, the Association of Finnish Regional and Local Authorities delegated Qualisan Oy to administer the use of the RAFAELA system. Qualisan Oy steered and controlled its use in the national nursing intensity benchmarking of hospitals in Finland. It set some criteria that wards had to fulfil to be allowed to take part in the national benchmarking of nursing care. Their nurses, for example, had to participate in the system's formal education, the system must have been in use in the ward for at least six months, and the inter-rater reliability had to be at least 70%. This indicated that at least 70% of parallel classifications of patients had to result in the same nursing intensity class. The OPC became OPCq to differentiate the uniform national version of it from the earlier local versions.

Under the coordination of Qualisan Oy, hospitals have been developing new measures, for example, for outpatient (POLIHOIq), psychiatric (PPC) and perioperative use. In summer

2007, the administration of the RAFAELA system was transferred from Qualisan Oy to FCG Efeko Oy. By June 2008, the use of the RAFAELA system had spread to about twenty hospitals and about 420 wards within specialized health care in Finland.

5.4 The previous validation process of the RAFAELA system

5.4.1 The previous validation process of the OPC

In addition to the validation being made in this thesis, the validity of the OPC had previously been analyzed in several studies in Oulu and Vaasa (Kaustinen 1995, Fagerström *et al*. 1998, Fagerström *et al*. 1999, Fagerström 2000). The main results of these are presented in Table 6.

Table 6. The previous validation process of the OPC

Type of validity	**Method**	**Closer description of used the method**	**Result**	**Ref.**
content validity	expert panel of nurses	evaluation in two different stages by expert groups representing staff nurses, nursing research in different universities and hospital administration at Oulu University Hospital	relatively high content validity	Kaustinen (1995)
content validity	expert panel of nurses	expert groups comprising a total of 68 nurses at Vaasa Central Hospital in 1998	relatively high content validity, but some weaknesses (see text)	Fagerström 2000
content validity	patient per-spective	75 patients' experiences of their own caring needs compared with the content of the OPC at Vaasa Central Hospital	patients' existential needs did not emerge clearly enough, and the instrument should be supplemented by a caring perspective; the physical and psychological needs were satisfactorily represented in the instrument	Fagerström *et al*. 1999
content validity	patient and nurse per-spective	a group of nurses' assessments of 73 inpatients' caring needs compared with the patients' own perceived caring needs at Vaasa Central Hospital	from the patient perspective, the OPC offered possibilities of providing an overall picture of the patients' nursing intensity and was, therefore, able to serve as a reliable basis for decisions concerning staff planning	Fagerström *et al*. 1998

In the validity study made at Vaasa Central Hospital (Fagerström 2000), it was considered that the content validity of the OPC method was relatively high, but there were, however, some weaknesses that the expert panel thought should be attended to, in particular, these were the content of subclasses of needs 1 (planning and co-ordination of nursing care) and 6 (teaching, guidance in care and follow-up care, emotional support) in the basic manual. The psychological and spiritual needs, as well as the nursing care activities connected with these, needed to be clarified. It was concluded that the OPC instrument and the whole RAFAELA system constituted an efficient and plausible instrument for the HRM of nurses and nursing care and at the same time, was a method by means of which the content of nursing care could be developed. (Fagerström 2000).

5.4.2 The previous validation process of the PAONCIL

The validity of the PAONCIL has been analyzed in two studies prior the sub-studies of this thesis, see Table 7.

Table 7. The previous validation process of the PAONCIL

Type of validity	**Method**	**Closer description of used the method**	**Result**	**Ref.**
criterion validity	association between the OPC and the PAONCIL	the association of the PAONCIL with the OPC as independent variable, analyzed by means of linear regression analysis in ten wards at Vaasa Central Hospital	the explanatory power varied from 19% to 59%; interest tended to decrease over the two-month period	Fagerström *et al*. 1999
content validity	methodological triangulation: questionnaires and focus group interviews	124 nurses assessed the extent to which the seven levels of nursing intensity of the guidelines for the PAONCIL corresponded to the nurses' experience of nursing intensity levels	the validity of the PAONCIL method was assessed as acceptable and the content of the PAONCIL manual was considered as being on a general level and not too differentiated for different specialties	Fagerström *et al*. 2002

5.5 The previous reliability testing of the RAFAELA system

Kaustinen (1995) evaluated the inter-rater reliability of the OPC at Oulu University Hospital. The result of the reliability evaluation showed an agreement of about 48-68% and a weighted agreement ranging from 87 to ninety-six per cent. The reliability of OPC was tested at eight wards concerned by means of inter-rater classifications (a total of 1080 parallel classifications) (Fagerström & Rainio 1998) The average agreement percentage was 75% and varied from 68% to 85%, which can be regarded as a good result.

The reliability of the OPCq-instrument is currently tested annually in the national benchmarking system under the administration of FCG Efeko Oy, and an agreement percentage of 70% is the prerequisite for participation in benchmarking.

6 The aims and research questions of the study

6.1 Aims

There are numerous patient classification systems in the world. Many of them play a central role in the allocation of nurse resources. Analyses of patients' care needs and the quality of patient care can be based on PCSs. Moreover, important economic calculations and plans dealing, for example, with budgeting, costing out nursing services, billing and cost control are made based on their results. Thus, there are numerous PCSs and they are widely used for many purposes – but are their results reliable? Only relatively few of the existing PCSs have undergone a proper scientific evaluation. Moreover, the validity and feasibility of patient classification systems is context-specific; the organization, practices and culture of health care and nursing care may have an influence on results. Therefore, the results gained, for example, in the US cannot automatically be generalized to health care in other countries, for example, Finland.

The RAFAELA system has become widely and ever increasingly used in Finland. It has, therefore, become important to become convinced of its properties; that they indeed allow such a wide and multifaceted use. Although scientific evidence on its validity and reliability already existed, it was considered important to strengthen further this kind of scientific basis of the RAFAELA PCS. Thus, at a general level, the aims of this study were to evaluate whether the RAFAELA system is valid and feasible enough to be used as a measurement tool for human resource management in nursing in the wards of Finnish specialized health care.

6.2 Research questions

The seven research questions studied are presented at a general level in Table 8. The results are presented so that all the study questions are considered one by one. The first five issues are presented in the same chapter, because they all deal with the same topic: testing the properties of the RAFAELA system and its components. Detailed analyses of these five issues and also of the sixth issue have been performed, and their complete results are presented in the publications of those sub-studies. The sixth and seventh issues are presented separately,

because they deal with separate entities. Moreover, the last study question also differs from the others in that it is more comprehensive. It is based on all the five sub-studies and the literature reviews of this thesis.

Table 8. The general issues that are studied

Issue	**Sub-study**
1. The determination of optimal nursing intensity	the first sub-study
2. Analysis of the properties of the PAONCIL score	the first sub-study the fourth sub-study the fifth sub-study
3. Analysis of the validity of the OPC measure	the second sub-study
4. The prerequisites for a successful analysis of optimal nursing intensity	the first sub-study the third sub-study
5. The role of non-patient factors in RAFAELA	the fourth sub-study
6. The association of work overload and increased sickness absenteeism among nurses	the fifth sub-study
7. The role of the RAFAELA system as a measurement tool for HRM in nursing	all sub-studies and this thesis

.

At sub-study level, the aims were operationalized to the following study items in the five sub-studies:

The first sub-study (*Professional Assessment of Optimal Nursing Care Intensity Level: A New Method for Resource Allocation as an Alternative to Classical Time Studies*.).

To ascertain:

1. if it was possible to determine the optimal nursing intensity of various wards, expressed in OPC points per nurse, using the PAONCIL instrument;
2. what this optimum was;

3. how it varied from ward to ward;
4. what the tolerance interval around this optimal nursing intensity was.

The second sub-study (*Validation of a new method for patient classification, the Oulu Patient Classification.*)

To test:

1. the concurrent validity of the OPC if PAONCIL was used as a 'gold standard';
2. the construct validity of the OPC by analyzing the relative importance (weight coefficients) of all the six sub-areas;
3. the construct validity of the OPC by analyzing the relative importance (weight coefficients) of the four nursing intensity categories;
4. the construct validity of the OPC by analyzing if factors such as age, gender and patient turnover had an independent explanatory power on nursing intensity or whether these were mediated through the OPC.

The third sub-study (*Determining optimal nursing intensity: the RAFAELA method.*)

1. To identify the minimum requirements for determining optimal nursing intensity that allow the results to be accepted as correct, in terms of:

- the length of the PAONCIL examination period;
- the PAONCIL questionnaire response rate;
- the explanatory power of the regression analysis;
- the mean values of the OPC and PAONCIL instruments.

2. To determine whether a standard value of optimal nursing intensity, 'a gold standard', existed for adult wards.

The fourth sub-study: (***Are nurses' assessments of their workload affected by non-patient factors? An analysis of the RAFAELA system.***)

To ascertain:

1. the contribution of non-patient factors to the RAFAELA patient classification system (non-patient factors mean factors other than nursing intensity that might have affected the workload of nurses, for example, administration, co-operation and so forth)
2. the explanatory power of nursing intensity and non-patient factors with regard to PAONCIL workload.

The fifth sub-study (*What degree of work overload is likely to cause increased sickness absenteeism among nurses? Evidence from the RAFAELA patient classification system*).

1. To examine whether nurses' work overload, as assessed by the RAFAELA system, was associated with increased risk of sickness absence.
2. To quantify potential loss of working days resulting from excess sickness absence related to work overload.

7 Materials and methods

The materials and methods of the five sub-studies are presented in detail in the following two chapters. Their central characteristics have been gathered and are presented in Table 9.

Table 9. The study objects, data and methods of the five sub-studies.

Sub-study	Study objects	Data	Methods
1) Professional Assessment of Optimal Nursing Care Intensity Level: A New Method for Resource Allocation as an Alternative to Classical Time Studies.	eight somatic wards and 148 nurses at one hospital in 1996-1997	daily OPC results (n = 19 324) and daily PAONCIL results (n = 8 458)	simple linear regression analysis
2) Validation of a new method for patient classification, the Oulu Patient Classification.	as above	as above	simple linear regression analysis, polynomial regression analysis, analysis of variance
3) Determining optimal nursing intensity: the RAFAELA method.	61 wards from eight Finnish secondary health care hospitals in 1997-2001	the optimal nursing intensity analyses of each ward: OPC and PAONCIL values, results of the analysis of optimal nursing intensity and the specialty of the wards	simple linear regression analysis, Spearman's rank correlation analysis, one-way analysis of variance, t-test
4) Are nurses' assessments of their workload affected by non-patient factors? An analysis of the RAFAELA system.	4 870 questionnaires in 22 somatic wards of a secondary health care hospital in 2002	OPC, PAONCIL and additional questions	Chi-square test, Spearman's non-parametric and Pearson's correlation analysis, (multiple) linear regression analysis and factor analysis
5) What degree of work overload is likely to cause increased sickness absenteeism among nurses? Evidence from the RAFAELA patient classification system	5 Finnish hospitals: 31 wards and 877 nurse employees in 2004	nurses' workload as a ratio of observed nursing intensity per nurse divided by optimal NCI per nurse, sick leave and demographic data	Negative binomial regression analysis

7.1 Materials

The first and second sub-studies comprised eight somatic wards and 148 nurses at the Vaasa Central Hospital, of whom about 80% were registered nurses and the remainder, 20%, were practical nurses. The period of investigation comprised about 3 months for each ward between October 1996 and February 1997. Data concerning the daily OPC results (n = 19 324) of all patients and the daily PAONCIL results (n = 8 458) of all nurses and daily nurse resources were collected.

In the third sub-study, 61 wards from eight Finnish secondary health care hospitals were included. All of these were wards subjected to optimal nursing intensity analysis between 1997 and 2001, for which sufficient data were available. All the wards used the RAFAELA system routinely. The following data were collected from the optimal nursing intensity analyses of each ward: OPC and PAONCIL values, results of the analysis of optimal nursing intensity (optimal nursing intensity, prevailing nursing intensity mean ± SD, daily PAONCIL averages' mean ± SD, duration of examination, response rate, possible outlier days in performed analyses) and the specialty of the wards.

In the fourth sub-study, existing data on the OPC, PAONCIL and additional – or non-patient – questions (Table 10) were used. These questions were developed to discover factors other than nursing intensity that might have affected the workload of nurses. Data were collected from 22 somatic wards of a secondary health care hospital on the west coast of Finland. This material was originally collected between September 29th and November 12th, 2002 for routine analysis of optimal nursing intensity. The length of the research in the wards was on average 25 days (range 21- 42 days). Altogether 4 870 questionnaires were returned. The response rate was 89.5 per cent.

Table 10. The list of additional non-patient issues studied with the RAFAELA patient classification system.

Non-patient issues
1. Managerial planning and organization of the work
2. Planning of the work rota
3. Substitute situation
4. Meetings, training
5. Students
6. Co-operation with physicians
7. Co-operation/co-ordination with other staff groups
8. Co-operation within the organization, e.g. different units
9. Co-operation in your own group
10. My own working ability (tiredness, common cold, worries)
11. Mental stress (terminal treatment, resuscitation)
12. Other factors

In the fifth sub-study, data on both the workload and sick leave of nurses were available from five Finnish hospitals: 31 wards and 877 nurse employees (range 13-54 per ward). The workload data from the first half of the year 2004 and sick leave data for the whole of the year 2004 were used. The average nursing intensity per nurse in the ward during the study period of January 1st to June 30th, 2004 in the Finnish nursing intensity benchmarking data were divided by the optimal nursing intensity per nurse. This ratio was used as a workload score, a measure of nurses' workload optimality in different wards. Data on sick leaves as short (1-3 days) and long (>3 days) spells of sickness absence were collected from hospitals' electronic records by the Finnish Institute of Occupational Health.

"Ethical guidelines for nursing research in the Nordic countries" published by the Northern Nurses Federation (1995) provided the guidelines for the planning and implementation of the empirical studies. In the participating hospitals, the patients' nursing care intensity was measured routinely on a daily basis with the OPC instrument. To protect confidentiality, the data the hospitals sent directly to researchers for the OPC and the PAONCIL analyses and the benchmarking data produced by Qualisan Oy did not contain any information about the

identity of the patients and nurses. Thus, the data were treated entirely anonymously. Each hospital gave the researchers official permission to use the data for scientific purposes. Approval of the ethics committee of the Finnish Institute of Occupational Health was also obtained to study the sickness absence records of the employees.

7.2 Methods

All sub-studies were observational (non-experimental) and retrospective. Various versions of the statistical program package SPSS for Windows were used to analyze the material. Some basic analyses were made with MS Excel. The principal statistical methods were various types of regression analyses that were employed in each of the sub-studies. The prerequisites of regression analysis were always analyzed and fulfilled.

In the first and second sub-studies, the main idea of the analysis was to use the average nursing intensity score per nurse (OPC value per nurse; independent variable) to predict the average PAONCIL value for the same calendar day. The material was analyzed by means of simple linear regression analysis. The average PAONCIL value for each calendar day was used as a dependent variable. Thus, by means of a regression equation, it was possible to calculate the optimal nursing intensity points per nurse, that is, the value which led to the average PAONCIL value zero. This constituted the optimal nursing intensity point. In the second sub-study, data were also analyzed by polynomial regression analyses. Analysis of variance was also used.

In the third sub-study, a possible 'gold standard' for optimal nursing intensity was determined graphically and on the basis of the parameters of distribution. Linear regression analysis was used to determine the minimum length of the PAONCIL examination period required. The mean of the OPC and PAONCIL values and the importance of the distribution of their values, measured as their standard deviation, were examined by means of Spearman's rank correlation analysis. The smallest sufficient response rate and explanatory power, as well as the relation between the variables of the mean PAONCIL and the optimal nursing intensity per nurse were determined by one-way analysis of variance. The smallest sufficient response rate was also analysed with two independent-sample t-tests.

In the fourth sub-study, the answers to the non-patient questions were first coded as follows: “added to workload” = +1,”reduced workload” = -1, empty = 0. Thus, “no effect”, “cannot say” and “not answered” all became coded as zero. The relation between answering activeness and time was analyzed using Spearman’s non-parametric correlation. When studying the association between the non-patient questions, the daily average of different wards was used as data. These data were analyzed using Pearson’s correlation analysis and factor analysis. Principal Axis Factoring was used as an extraction method.

The relations between the daily averages of the wards of these variables were analyzed using Pearson’s correlation coefficient and a multiple linear regression analysis. In the regression analysis, PAONCIL values were used as a dependent variable and OPC-points/nurse and the value of the six non-patient factors (factors other than nursing intensity that might affect nurses’ experience of their workload) with independent explanatory qualities were used as independent variables. The stepwise backward regression method was used.

Daily averages of different wards were used in the correlation analyses and most regression analyses instead of the results of individual questionnaires. This choice was made so that it would be possible to use similar data to that which is routinely used for determining a ward’s optimal nursing intensity. The second reason was that an answer to one individual non-patient question had only been given in under 4% of the questionnaires. If “no affect” or “no answer” had been coded as 0 and the answers as +1 and -1 there would have been a falsely high correlation between the additional questions, resulting from the varied level of answering activity between nurses.

In the fifth sub-study, the association of work overload and sick leaves was analyzed. Workload was expressed as a workload score, so, the actual average workload (OPC nursing intensity points per nurse) divided by the optimal nursing intensity per nurse. This workload score was analyzed both as a continuous and as a categorical variable. In the latter, the workload score was divided into four categories, using 1.00, 1.15 and 1.30 as cut-off points. The cut-off point of 1.00 is a workload corresponding to the optimal value and 1.15 corresponding to the upper limit of the optimal interval: 15% above the optimal value. This categorization was used to achieve categories with sufficient power to detect moderate sick leave differences. For each individual, the numbers of short and long spells of sickness

absence were computed, and the follow-up period was measured in person-months. Rates of sickness absence spells were computed and expressed per 10 person-years.

Negative binomial regression models were applied in the analyses because traditional regression models, which are based on the normal distribution, may lead to serious problems in modelling a dependent variable following a negative binomial distribution. Separate regression models for self-certified and medically certified sick leave were used. Rate ratios and their 95% confidence intervals of short and long spells of sickness absence were calculated for univariate analyses and they were then adjusted for age and gender in multivariate analyses.

8 Results

8.1 Testing of the validity and other properties of the RAFAELA system and its components

8.1.1 The determination of the optimal nursing intensity with the RAFAELA system

The determination of optimal nursing intensity could be performed successfully (the criterion for this was a determination coefficient of at least 0.25) in five out of eight wards for adults in the first sub-study. The average determination coefficient in the various wards was 0.37. Thus, the PAONCIL instrument explained 37% of the variation of nursing intensity per nurse. This corresponds to a linear correlation of 0.60. The optimal nursing intensity scores per nurse were close to each other in five out of six wards for adults, ranging between 3.0 and 3.6 'class coefficient points' per nurse, corresponding to 20.2 – 24.3 OPC points per nurse. Figure 5 on page 71 shows a typical scattergram of linear regression analysis.

8.1.2 Analysis of the properties of the PAONCIL score

In the first and fourth sub-studies, the distribution of the PAONCIL scores was similar to the normal distribution, but the 0-value was to some extent over-represented (Figure 6). In the second sub-study, in separate univariate analyses, the PAONCIL value was independent of the amount of nurses' working years and of different occupational categories.

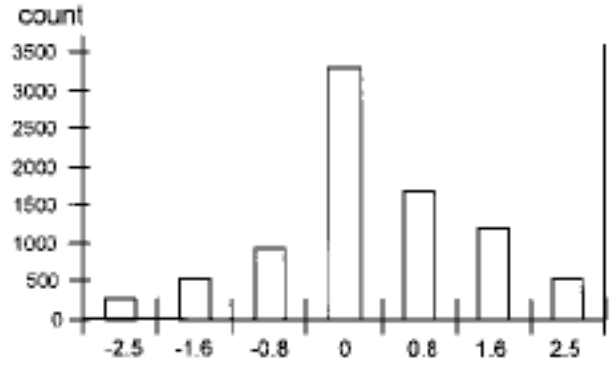

Figure 6. Distribution of the values of PAONCIL.

The classes are -2.5 ± 0.5, -1.6 ± 0.4, -0.8 ± 0.4, 0.0 ± 0.4, 0.8 ± 0.4, 1.6 ± 0.4 and 2.5 ± 0.5. The figure is from the first sub-study and is published with permission of Scand J Caring Sci and Wiley-Blackwell Publishing.

8.1.3 Analysis of the validity of the OPC measure

In the second sub-study, the determination coefficient, that is, the explanatory ratio of the linear regression analysis, did not increase if the following variables were added to the regression equation as independent variables: the number of patients, the average age of the patients and the distribution of gender. Thus, the OPC measure took into consideration those variables. This supported its construct validity. As a further test of construct validity, an attempt was made to determine the adequacy of the weighing coefficients of the six sub-areas and four levels of nursing care. Because of the size of the material, the distribution of its parameters and the clear sample-based multicollinearity of the independent variables, this proved to be impossible.

In the determination of optimal nursing intensity, the determination coefficient was found to be 0.36, indicating a fairly good association between the OPC and PAONCIL measures, considered as a ‘gold standard’ in this comparison. Thus, this could be considered as testing of concurrent validity.

The performance of the OPC did not clearly deteriorate even if, for the sake of clarity and concreteness, summing up the whole ward's nursing intensity points was abandoned and every patient was first, based on these points, divided into one of four nursing intensity categories, each of which with its own weighting coefficient. The nursing intensity of the whole ward was, in this praxis, then calculated summing up these weighting coefficients of every patient. This praxis dominated in the early years of the use of the RAFAELA, but was later abandoned as unnecessary.

8.1.4 The prerequisites for a successful analysis of optimal nursing intensity

Three out of eight wards in the first sub-study had an insufficient determination coefficient. They were found to be wards with a low daily variation in nursing intensity per nurse. It was speculated that insufficient information to the nurses constituted another cause. This topic of the assumptions that must be fulfilled in order to be able to rely on the results obtained was analyzed more thoroughly in the third sub-study.

An analysis of **the smallest sufficient explanatory power** was performed. An explanatory power of about 0.25 (or 25%) was found to be the smallest undoubtedly sufficient explanatory power. This smallest sufficient explanatory power was shown both visually and by using one-way analysis of variance. Visually this was shown by inspection of a scattergram, the horizontal axis of which was the explanatory power of the performed optimal nursing intensity analysis, and the vertical axis the optimal nursing intensity per nurse of the same analysis. Thus, each dot represents the analysis results of one ward. Specialties with known unusual optimal nursing intensity (paediatrics, neurology, rehabilitation) (Fagerström & Rauhala 2003) were excluded from the data. It can be seen (Figure 7) that the points of optimal nursing intensity assume a funnel-shaped formation: when the explanatory power increases, they are distributed over an increasingly narrower zone until, with the explanatory power exceeding 0.25 (or 25%), the distribution no longer becomes narrower. The subgroup of 32 wards with an explanatory power exceeding 25% was analysed further. It was divided into three groups of approximately equal size: explanatory power 25% – 34.1%, 34.2% – 41% and above 41%. Using one-way analysis of variance, there were no statistically significant differences between the three groups ($p = 0.429$).

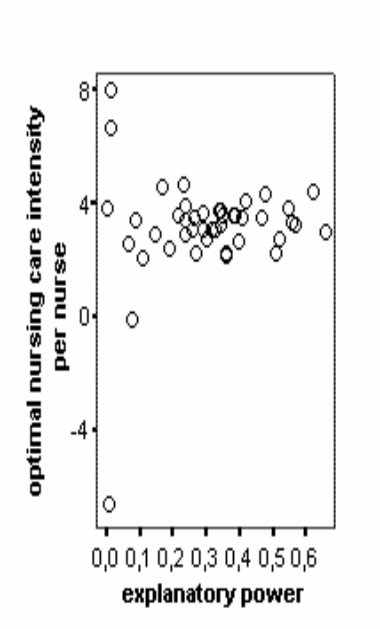

Figure 7. Relation between explanatory power and point of optimal nursing intensity (in weight coefficient points), a scattergram from the optimal nursing intensity analyses of 48 wards.

Specialties with an unusual optimal nursing intensity have been excluded from the data. This figure is from the third sub-study and is published with permission of J Adv Nurs and Wiley-Blackwell Publishing.

The existence of a potential 'gold standard' of optimal nursing intensity was also studied. No 'gold standard' of optimal nursing intensity, common to all adult wards could be found (Figure 8).

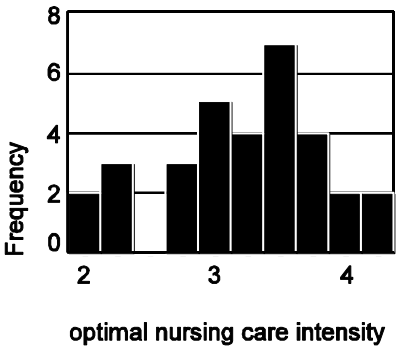

Figure 8. Histogram illustrating the distribution of optimal nursing intensity per nurse (in weight coefficient points) in 32 wards.

The specialties of paediatrics, neurology and rehabilitation are not included in the analysis, nor are wards with an explanatory power of under 25%. No 'gold standard' of optimal nursing intensity could be found. This figure is from the third sub-study and is published with permission of J Adv Nurs and Wiley-Blackwell Publishing.

The sufficient duration of the PAONCIL analysis period was determined. The values of optimal nursing intensity were not noticeably affected after 2.5 – 3 weeks (Figure 9) and the explanatory power hardly changed after a month of follow-up. Thus, 3-4 weeks can be considered as a sufficiently long follow-up period.

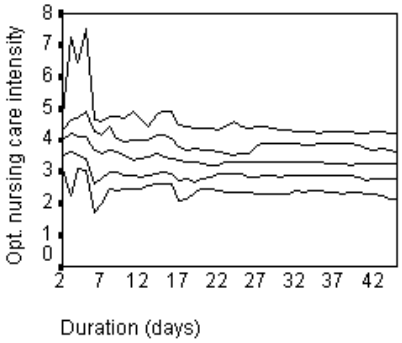

Figure 9. Optimal nursing intensity as a function of the duration of OPC and PAONCIL measurements (days).

The graphs from top to bottom are 90, 75, 50, 25 and 10% percentiles of the distribution of optimal nursing intensity per nurse (in weight coefficient points) in different wards.

Specialties with a deviating optimal nursing intensity are excluded from the data (n = 29). This figure is from the third sub-study and is published with permission of J Adv Nurs and Wiley-Blackwell Publishing.

The sufficient response rate in the PAONCIL analysis was then analyzed, using one-way analysis of variance. The data consisted of 51 wards. The specialties with a deviating optimal nursing intensity were excluded. The wards were divided into five groups of approximately equal size: response rate below 70%, 70 – 75%, 75 – 80%, 80 – 85% and above 85%. The association between the response rate and the values of optimal nursing intensity is shown in Figure 10. With a response rate of 70 – 75%, optimal nursing intensity values were at the same level as with higher response rates. In the analysis of variance, no statistically significant differences were found in the values of optimal nursing intensity between the five groups (p = 0.191). However, there were so few wards with a response rate below 70% that it could be shown only that a response rate above 70% is at least sufficient.

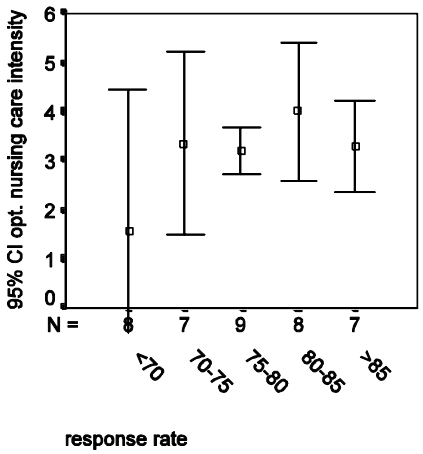

Figure 10. The relation between the response rate of the PAONCIL period and optimal nursing intensity.

The wards (n = 39) were divided into five groups according to the response rate. This figure is from the third sub-study and is published with permission of J Adv Nurs and Wiley-Blackwell Publishing.

The significance of the mean PAONCIL value.

It was found that the value of the optimal nursing intensity per nurse remained the same even when the mean PAONCIL value rose as high as 0.65, after which it began to decrease. This was shown both visually (Figure 11) and by one-way analysis of variance. Thus, the average PAONCIL value of 0.65 can be considered as the upper limit of reliable analyses of optimal nursing intensity.

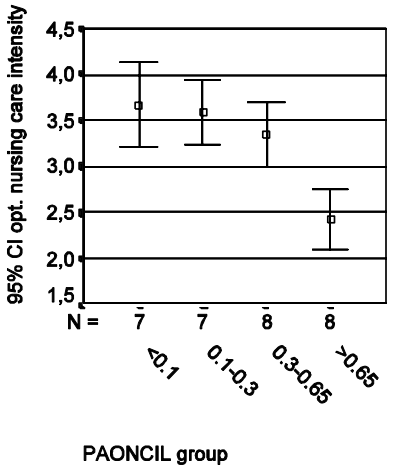

Figure 11. The relation between the mean PAONCIL value and optimal nursing intensity.

The PAONCIL values are divided into four categories. The graph presents the mean and the 95 per cent confidence intervals of the optimal nursing intensity in wards with various average PAONCIL values. This figure is from the third sub-study and is published with permission of J Adv Nurs and Wiley-Blackwell Publishing.

8.1.5 The influence of non-patient factors on the RAFAELA system

The influence of non-patient factors (factors other than nursing intensity that might affect nurses' experience of their workload, analyzed by the PAONCIL) on the RAFAELA patient classification system was studied in the fourth sub-study. Response activity was as follows: at

least one non-patient question was answered in 1 288 questionnaires out of 4 870 (26.4% of the questionnaires). Those who answered these questions ticked 1.8 answers on average.

The associations between non-patient questions were analyzed using the daily average of the answers. In the correlation analysis, questions (Q1 to Q5) associated with management, organization and staffing clearly correlated positively with each other, as in turn did questions (Q6 to Q9) associated with co-operation. In the factor analysis, the original factors with the highest eigenvalue (>1) could explain 54.2% of the variation and after the rotation, a third (34.4%). Eight questions out of twelve were grouped into four groups:

- administration (Q1 and Q2)
- staff and time resources (Q3 and Q4) and mental stress (Q11)
- co-operation within units (Q7 and Q9)
- co-operation between units. (Q8)

In the analysis of **association between different non-patient questions and workload**, the following non-patient factors had a statistically significant Pearson's correlation with workload, measured by PAONCIL: meetings (r=0.16), co-operation with doctors (r=0.11), with other staff groups (r=0.10) and within the entire organization (r=0.17), mental stress (r=0.19) and other factors (r=0.35).

The impact of non-patient factors on the explanatory value of the OPC to the variation of the PAONCIL was also analyzed. A usual analysis of optimal nursing intensity was performed in all of these 22 wards, using linear regression analysis. The daily OPC nursing intensity points per nurse for that ward were considered as an independent variable and the PAONCIL average of the corresponding ward as a dependent variable. In this analysis the median of the explanatory value was 0.45 (45%) and a sufficient explanatory value, over 25%, was achieved in 17 out of 22 wards. Next, the above-mentioned six non-patient factors with independent explanatory qualities were also included as independent variables in the previous regression analyses, using the stepwise backward regression method. The PAONCIL average of the same ward on the same day was still the dependent variable. When these additional factors were taken into the model as well as the OPC nursing intensity, the model's median of the explanatory value rose from 0.450 (45.0%) to 0.548 (54.8%).The following variables were left in the model as independent explanatory variables in at least some wards:

- OPC (15 out of 22 wards)
- "other factors" (Q12; five wards)
- "mental stress" (Q11; three wards)
- "Co-operation/co-ordination with other staff groups" (Q7; two wards)
- "Meetings, training" (Q4; one ward)
- "Co-operation with doctors" (Q6; one ward)

8.2 The association of work overload and increased sickness absenteeism among nurses

The fifth sub-study studied the association between nurses' work overload, indicated by nursing intensity per nurse in relation to the optimum, and their sick leaves. The mean workload was 9% above the optimal value, with a standard deviation of 18%. There were 27.0 sick leaves per 10 person-years. Two-thirds of these were short, self-certified absence spells and one-third were medically certified spells of more than 3 days (Figure 12).

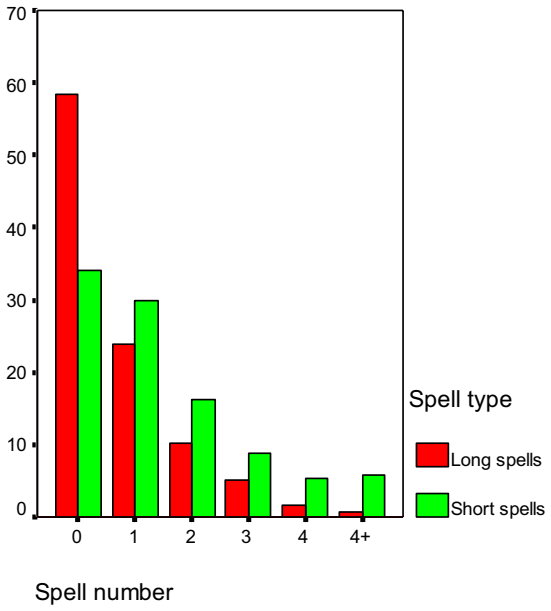

Figure 12. Distribution of the number of nurses' short (1 to 3 days) and long (longer than 3 days) spells of sick leave per person-year.

This figure is from the fifth sub-study and is published with permission of J Adv Nurs and Wiley-Blackwell Publishing.

In negative binomial regression models adjusted for age and sex, the rates of self-certified and medically certified sick leaves were 1.44 (95% CI 1.13-1.83) to 1.49 (1.10-2.03) times higher among nurses with a workload ≥30% above the optimum than among those with an optimum workload. Among nurses with a workload 15-30% above the optimum, the corresponding rates of self-certified and medically certified sick leaves were 1.45 (95% CI 1.22-1.74) and 1.28 (1.01-1.63) times higher than among those with an optimum workload (Figure 13). The results were quite similar in the univariate analysis.

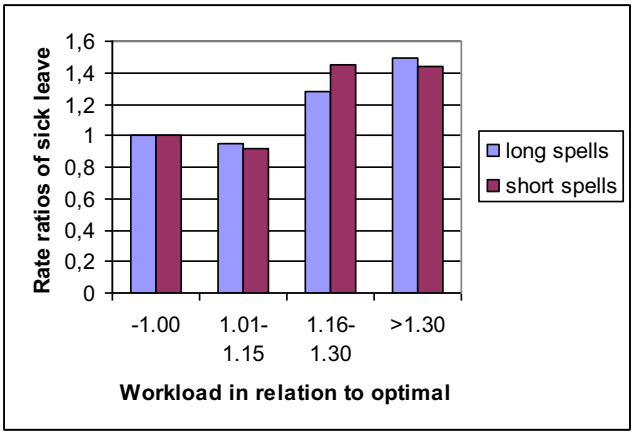

Figure 13. Rate ratios of self-certified sickness absence (spells of 1 to 3 days) and medically certified sickness absence (spells longer than 3 days) by workload in relation to optimal value (n=877).

Multivariate models. Other categories in this model are always compared to a reference category, given the value of 1.00.

The corresponding excess rate of sickness absence was 12 days per person-year, when the workload exceeded the optimum by more than 30% (Figure 14).

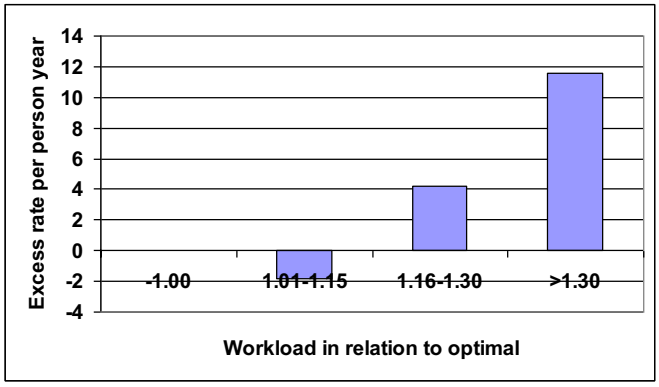

Figure 14. Excess sickness absence days per person-years in relation to workload optimality among Finnish nurses (n=877).

8.3 The role of the RAFAELA system as a measurement tool for HRM in nursing

In the following, the principles of HRM and the possible role of the RAFAELA within it are studied, based on the literature reviews of this thesis and the results of the sub-studies. The question of how these roles are realizable and actually come true in the everyday use of these systems needs studies of its own. In the following, this topic is analyzed more from a theoretical basis.

The central principles of HRM can be listed in many ways, one proposal, based mainly on Tables 3 (Storey 1989 and Vuori 2005) and 4 (Storey 1995) is presented in Table 11. These principles are discussed in the text one by one.

Table 11. The central principles of HRM and the possible theoretical role of the RAFAELA system in this context.

Principle of HRM	**The possible role of the RAFAELA**
Job satisfaction together with maximal performance of HR	++
It is the human resource which provides the competitive edge; therefore, HRM decisions are of strategic importance	±
Management involvement is necessary	++
Management culture is more important than managing procedures and systems	±
Integrated action on selection, communication, training, reward and development	–
Recruitment is important, employees should be very carefully selected	–
Increasing employees' influence, self-direction and control is stressed	+
Performance, not hierarchical position, determines the reward	±
Individual work contracts	–
Working conditions are in harmony with the worker's individual needs	+
Restructuring and job redesign to allow devolved responsibility and empowerment.	++
Division of labour is based on teams	–
Employees should be very carefully developed; promoting learning and training is essential	+
Solutions to conflicts between individuals are sought from the organizational climate and culture	±
Flexibility is emphasized in following rules	+
Information and communication policy promote employees' self direction	+
The main aim should be not mere compliance with rules, but employee commitment	+
The importance of measuring	+++

Explanations of the symbols: +++ = strong role, ++ = relatively strong role, + = minor role, ± = uncertain role, – = no role. The list of principles of HRM modified mainly from Tables 3 (Storey 1989 and Vuori 2005) and 4 (Storey 1995).

Job satisfaction together with maximal performance of human resources – As a system for measurement, the RAFAELA, of course, does not commit itself on this item as to whether the workload should be maximal or optimal. However, it creates possibilities for the improved monitoring of the optimality of nurses' workload. With satisfied, committed and capable workers, that full capacity can, in principle, also become higher. By using this system, it has also been shown (sub-study 5) that work overload is associated with adverse consequences that managers should be aware of.

It is the human resource which provides the competitive edge, therefore, HRM decisions are of strategic importance – The RAFAELA does not influence the importance of HRM decisions, but, in principle, it can contribute to the precision of them by providing information on nurses' workload and additional information on its costs.

Management involvement is necessary – The RAFAELA could, in principle, promote top and line management involvement by providing them with quantitative information of good quality on the nurses' workload and the case mix of patients.

Management culture is more important than managing procedures and systems – Of course, the RAFAELA does not take an attitude to valuing these items, but by increasing the possibilities of managers to distribute nursing work justly, it can promote employees' feeling of justice and thus, promote organizational culture.

Increasing employees' influence, self-direction and control is stressed – The RAFAELA system is a bottom-up model with nurses determining the values of both the OPC and the PAONCIL, showing appreciation of employees' professionalism and judgement and supporting employees' empowerment in theory by increasing their influence.

Performance, not hierarchical position, determines the reward – In the primary nursing model, individual nurses have 24-h total nursing responsibility for individual patients for their entire stay in the ward. In this model, using the RAFAELA, it is, in principle, possible to measure individual nurses' nursing intensity, workload and productivity. Thus, the RAFAELA enables individual rewarding in this setting. Similarly, in team nursing, it enables the determination of the team's workload, and productivity.

Working conditions are in harmony with the worker's individual needs – See also comments on the previous item. It is possible to quantify a ward's – or in team nursing, a team's – nurses' average workload, but only the individual workload is quantified in primary nursing.

Restructuring and job redesign to allow devolved responsibility and empowerment – By systemically utilizing the nursing intensity data from this bottom-up model system, it is possible to allocate nurse resources more justly. This could, in theory, increase the empowerment of nurses.

Employees should be very carefully developed; promoting learning and training is essential – The system can promote employees' co-operation and development and training in nursing care. The implementation and maintenance of this patient classification system needs training. It also automatically causes a need for discussion between the nurses of different wards and hospitals about the principles of classification – and of nursing routines. Differences between wards and hospitals may lead to official or unofficial comparisons between organizations. This benchmarking activity, in turn, may promote the development of nursing practices in a way that can also be called organizational learning.

Solutions to conflicts between individuals are sought from the organizational climate and culture – According to HRM, solutions to conflicts between individuals are sought from the organizational climate and culture. Experience of justice and perception of fairness in an organization promotes its culture and climate. When a valid PCS is used, principles of allocation are just and transparent and in harmony with the idea of optimal workload. This can, in principle, be thought to increase nurses' experiences of organizational justice or, at least, procedural justice, and in this way also promote employee commitment and organizational culture. The RAFAELA system is also a bottom-up model, with nurses determining the values of both measures, showing appreciation to employees' professionalism and judgement and supporting employees' empowerment and thus improving organizational climate and nurses' commitment. So, it can at least be thought that the utilization of this PCS promotes organizational culture.

Flexibility is emphasized in following rules – Nursing resources can be moved more flexibly according to need, based on nursing intensity results, instead of following 'rules' for the number of nurses that every ward has.

Information and communication policy promote employees' self direction – The RAFAELA system provides leaders and workers with more precise information, including benchmarking results. These can, in principle, promote nurses' activity and empowerment.

The importance of measuring; "You cannot manage what you cannot measure" – Nurses' workload is a central object of measurement. The validity of this PCS has received additional strength from the sub-studies of this thesis. Without a valid and creditable measurement system, it is impossible to conduct evidence-based health care administration and a just allocation of resources. Also, by yielding information on nurses' workload, it is easier to evoke top management involvement. Nursing intensity information should be part of a hospital's management support system.

9 Discussion

9.1 The validity, creditability and limitations of this study

This study consists of five sub-studies. Issues concerning validity, reliability and limitations of the sub-studies are discussed in detail in every sub-study. Therefore, only the more general and central items are presented and discussed here.

The study design in all of the sub-studies was observational. Contrary to experimental studies, in observational studies, study objects are observed or certain outcomes are measured. No interventions are made to affect the outcome. The materials of the studies are not random samples, but they can be considered as convenience samples or instead of sample studies, the studies are more like total studies, because practically all the data available in Finland at the time of each sub-study was gathered and employed in them. The materials consisted of hospitals' routine nursing intensity assessments in production use and routine PAONCIL studies for the determination of optimal nursing intensity. As such, they have the strengths and weaknesses of such real world materials. However, the aims have been mainly quite pragmatic: to evaluate the validity and feasibility of a patient classification system in routine use, in the real world with 'random' nurses, not in artificial and ideal laboratory conditions, with specially trained nurses. Therefore, the routine use nature of the material can primarily be considered as a strength. It can also be said that a measure of nursing care intensity does not measure the whole workload of nurses. However, the majority of nurse's working hours are still allocated to direct and indirect nursing intervention work related to individual patients, and the workload analyses focused systematically on this patient associated workload. Moreover, only those activities and those nurses were included in the calculations of nursing care intensity that were related to direct or indirect patient care.

Most often, the statistical methods used were various types of regression analyses. The applicability of such analyses has been judged in every sub-study. Moreover, in every sub-study, the fulfilment of the general prerequisites for linear regression analysis has been examined.

The materials in the first and second sub-studies were from one hospital (Vaasa Central Hospital) and consisted of all the wards that, at that time, used the whole of the RAFAELA system. The materials in the fourth sub-study were from all wards of another large central hospital in Finland (Satakunta Central Hospital), which at that time, was one of the few hospitals that had already successfully implemented this PCS in the whole hospital. Sub-studies three and five used material from several Finnish hospitals, again, including the whole material available at that time. In a way, the third sub-study can also be seen partially as an extension of the one-hospital first and second sub-studies to several hospitals. All the materials represent typical wards, patients and nurses in Finnish specialized health care, without any known possibility of significant bias. The data gathered have been carefully checked for many kinds of errors. Furthermore, in the large Finnhoitoisuus project, comprising nine Finnish hospitals (Fagerström & Rauhala 2003), it was shown that Finnish hospitals were quite uniform in their nature. Thus, there are no well-grounded reasons to suspect the internal validity of the sub-studies. Likewise, the results should be generalizable (external validity) to the wards of the whole of Finnish specialized health care.

The results of the seventh study question, the role of the RAFAELA system as a measurement tool for HRM in nursing, are based on the literature reviews of this thesis and on the five sub-studies. The analysis is theoretical.

9.2 Discussion of the main results of the five sub-studies

At first it was shown that, as a whole, the determination of optimal nursing intensity could be performed successfully in most wards, with a satisfactory explanatory power, using the RAFAELA system and its measures, the OPC and the PAONCIL and data of daily nurse resources. Thus, as a whole, this classification system was shown to function satisfactorily in the real world, in the context of Finnish specialized health care wards.

Next, the innovative solution of replacing time studies with a simple and time-saving bottom-up approach, the PAONCIL measure, was studied more closely. It was shown to be a sufficiently valid measure, linking nursing intensity points to the real world. Thus it was shown that laborious time studies were not needed in routine use of a valid PCS. Following this, the validity of the OPC measure was studied more closely. Its construct validity was

shown to be sufficient. Similarly, as part of the whole system, its concurrent validity with the PAONCIL as a ‘gold standard’ was also shown to be sufficient.

Based on the above analyses, most often the RAFAELA system was shown to work rather well. Nevertheless, what were its limits, and how did it function when exposed to exceptional circumstances? Following this, an examination of what constituted the prerequisites for a successful analysis of optimal nursing intensity was performed. Clear ranges of the proper function of the PCS could be determined, for example, regarding the explanatory power of the regression analysis, length of the PAONCIL follow-up period, response rate and upper limit of the average PAONCIL value. These analyses showed in which circumstances the PCS was still valid.

The analysis of the significance of the non-patient questions, concerning factors unrelated to the nursing activities of patients, showed that adding these non-patient factors to the regression equation already including daily OPC per nurse as an independent and daily average PAONCIL as a dependent variable did not markedly increase the explanatory power. Thus, this result also supported that the OPC and the PAONCIL instruments measure quite the same phenomenon and do not markedly measure such items that are unrelated to nursing activities concerning patients. The questions regarding non-patient factors could also be said to offer the managers of nursing work valuable and feasible practical information on the functioning and problems of a ward. These non-patient questions are not included in the RAFAELA system itself, but have sometimes been used alongside the PAONCIL analysis for scientific or practical purposes.

After all these results had been gained, combined with available previous research, the validity of the whole RAFAELA system and its elements could be considered as having being thoroughly analysed and evaluated.

Next, this recently thoroughly validated RAFAELA system was used to study the association between nurses’ workload and sick leave. The analyses showed that there was a clear linear trend between increasing workload and increasing sick leave. Among nurses with a workload >30% above the optimum, the rate of self-certified periods of sick leave was 44% higher, and for medically certified sick leave, 49% higher than among those nurses with an optimum workload. These excess rates of sickness absence resulted in 12 extra sick leave days per

person-year. About 5-6% of the increased productivity resulting from work overload was directly lost in increased periods of sick leave. Bearing in mind that sick leave is also a risk marker of future disability pension and an independent predictor of mortality (Kivimäki *et al.* 2003a, Vahtera *et al.* 2004), the negative long-term consequences to employees and economy can be even more serious. In a recent study, Virtanen *et al.* (Internet reference No 3) were able to show that overcrowding in hospital wards is also a predictor of increased antidepressant treatment among hospital staff. Moreover, multiple studies have shown an association between nurse staffing and patients' mortality and a variety of non-mortality outcomes (Aiken *et al*. 2002, Needleman *et al*. 2002, Sermeus *et al*. 2007). These findings suggest that work overload may significantly contribute to health problems among employees and also lead to a considerable risk of adverse outcomes to patients.

9.3 Evaluation of the whole validation process of the RAFAELA system

There are numerous nurse demand methods, nursing workload measurement systems and patient classification systems (Arthur & James 1994), but only a minority of them have undergone a proper and thorough formal scientific validation process. However, the scientifically demonstrated validity of a measurement instrument should be the *condicio sine qua non* for their wide-scale use.

Creating, developing, validating, piloting and implementing a measurement instrument is a demanding and time-consuming process, where different phases follow each other – or intermingle with each other – and are repeated. Validation, for example, should be followed by re-validation if modifications have been made to the instrument or the environment and the context has been changed. The OPC – or OPCq – measure has been further developed from an already validated Canadian HSSG classification system. In principle, the content validity of a nursing intensity measure can be tested by using expert panels of nurses or other personnel groups. The other available alternative is to use the patient perspective. The OPC has been tested in several studies and in both ways.

The criterion validity of a patient classification system can be tested using either time studies/work sampling methods or another instrument as a 'gold standard'. Traditionally the

need for personnel resources has been calculated on the basis of time studies and activity analysis (Giovannetti 1979, Alward 1983, Rosenbaum *et al*. 1988, Giovannetti & Moore Johnson 1990, Phillips *et al*. 1992, Mayo & Van Slyck 1999). Such studies have been made also with the OPC (Onnela & Svenström 1998**)**. In the studies of this thesis, the PAONCIL has been used as a 'gold standard' when testing criterion – or more precisely concurrent – validity. It has been repeatedly shown that it is possible to measure nursing intensity – and its relation to an optimum – using the OPC and the PAONCIL instruments. The prerequisites for a successful analysis have been studied in many ways.

The construct validity of a nursing intensity measure can be tested either by adding new parameters to the measure or by seeking more optimal coefficients for the existing parameters and their values. Both have been done – or at least the attempt has been made to do so – in this thesis. The former showed that the OPC's construct was valid; the latter was not possible for statistical-mathematical reasons.

The content validity of the PAONCIL instrument can, in principle, be scientifically tested by interviewing nurses. This has been done. It has also been shown that the OPC and PAONCIL measure much the same phenomenon and have a moderately high coefficient of determination in regression analyses. The twelve non-patient factors, for example, could only add a small amount of explanation to the PAONCIL value and did not compromise the reliability and operational ability of the RAFAELA patient classification system. Thus, the different types of validity of the whole RAFAELA system have been systematically tested in a variety of ways, with satisfactory results. Similarly, the inter-rater reliability of the OPC is tested scientifically – and inter-rater reliability testing is part of the routine use of the system and national benchmarking activity of nursing intensity.

9.4 Comparison of the characteristics and feasibility of the RAFAELA system to the requirements of modern patient classification systems

It is said that traditional patient classification systems no longer contributed usefully to the HRM of nursing. Value from these systems can only be realized through major revisions and

integration with other systems, linking patient outcomes, cost of care and quality measurements together (Finnigan *et al*. 1993, Aiken *et al*. 2002, Needleman *et al*. 2002, Welton *et al*. 2006, Sermeus *et al*. 2007). The provocative question is often posed of whether it really matters how much work the nurse performs (the time required to accomplish tasks) if the outcome of the work is not measured. According to Malloch and Conovaloff (1999), third generation PCSs were developed for this reason, focusing on optimizing the patient-caregiver process for the outcomes of cost, health and caregiver satisfaction.

The RAFAELA system itself does not, in routine use, produce data on costs, outcomes or quality, but it is completely possible to calculate the cost of a nursing intensity point (Fagerström & Rauhala 2001, Fagerström & Rauhala 2003, Pusa 2007). It can also be considered as a matter of taste with modern data systems, whether all the data is included in the PCS or if the outcome and cost data are integrated with nursing intensity data just in the IT system of the organization. A module of its own within the RAFAELA system would, of course, promote the creation of uniform follow up parameters and benchmarking between wards and hospitals. However, it is not enough to measure the outcomes of patients only. Excessive workload may have consequences also for the nurses themselves (see the fifth sub-study) in the form of increased sick leaves and other detrimental consequences.

It can be said that, to some degree, the RAFAELA system also includes quality aspects. The descriptions of the seven nursing intensity levels of the PAONCIL instrument indirectly contain quality aspects. Thus, for example, too high a nursing intensity (above zero in the PAONCIL score) implies a situation, in which the nurses are unable to respond to the patients' caring needs, in other words, the nursing intensity level is too high. In such situations, for example, risk of mistakes may increase. However, a direct measure of the quality of nursing care is not included in the RAFAELA system.

It is also possible to classify nurses' tasks into different groups based on the degree of skill or competence needed in performing them. In this way, PCSs could monitor the existing and needed skill mix and the difference between these two among nurses. In the RAFAELA system, the analyses of nursing intensity include both registered nurses and practical nurses of various wards. Thus, in the analyses of nursing care intensity, the amount of care (patient-associated workload) is monitored, not the skill level at which the care is provided. Thus, the RAFAELA system does not make any adjustments for skill mix. In the context of the HRM of

Finnish specialized health care, this cannot be considered a major problem, because the skill mix in wards of public hospitals is quite uniform, for example, the mean (±SD) percentage of second level or licensed practical nurses or corresponding of all nurses in wards is 29% (± 11%) (Fagerström & Rauhala 2003). There is no variation of practical significance in this between Finnish public hospitals of specialized health care. In Finland, untrained employees do not take part in patient care. Skill mix can also be considered easier to monitor in task-oriented rather than dependency-driven methods of nursing workload measures. However, a lack of skill-mix monitoring can, to some degree, be considered a deficiency in the RAFAELA system.

For the sake of simplicity, the coefficients of the four levels of the six sub-areas of nursing care are scored to 1, 2, 3 and 4 points. The explanatory power of the OPC of the variation of the PAONCIL has been sufficient with this construct. However, it can be considered as obvious that it is possible to trim the OPC instrument further, by allowing all 24 weights of the measure to receive an individual value, based, for example, on time studies. In this way, it is possible to increase further the explanatory power of the OPC instrument. At present, it is also possible to have only four-fold variation in the nursing intensity between patients, from six to 24 nursing intensity points. With further trimming of the coefficients it would be possible to show better the existing wider variety of nursing intensity and case mix across patients.

The RAFAELA system is administered by FCG Efeko Oy. This arrangement allows the possibility of maintaining uniform coding instructions practices, standards, training and development of the system. It also makes it possible to have nursing benchmarking between hospitals.

9.5 The role of the RAFAELA system in the HRM of nurses

The introduction of this thesis stated the need for the measurent of nurses' workload using a reliable patient classification system in the context of nurses' HRM. The RAFAELA system has many features that support the principles of HRM and these are analysed in detail in Chapter 8.3 of the Results. The use of the RAFAELA system can be thought to support the central concept of HRM: high job satisfaction and high productivity simultaneously.

However, at its best, even if scientifically validated and feasible, the RAFAELA is just a system of measures. Therefore, it can support the wide concept of HRM in nursing only to a certain degree and only in certain sectors. In principle, the use of the bottom-up design RAFAELA patient classification system can, of course, predominantly support such aspects of HRM that are associated with workload measuring, planning of working conditions, increasing the involvement of managers and empowerment of nurses, increased feeling of justice and so forth.

According to HRM, the labour resource should be utilized to its full capacity and efficiency (Storey 1989). It tries to avoid ineffective underutilization of working capacity (suboptimal workload) and, on the other hand, short-sighted exploitation of labour. Thus, employees' working capacity should be used optimally, in harmony with their resources. HRM emphasizes the importance of harmony between working conditions and workers' needs. A necessary prerequisite for such harmony is that nurses' workload is optimal. This can be assured by using the RAFAELA system. HRM also stresses that workers' individual needs must be taken into consideration. It also suggests that performance also determines the reward. In principle, it is possible to use the RAFAELA system as a measurement of nurses' productivity (Pusa 2007). This productivity of nurses can, in principle, in certain conditions (primary nursing) also be measured individually, but to the best of my knowledge there are no such applications in use.

The optimality is determined in the RAFAELA system by using the PAONCIL instrument. In principle, optimality could also be determined based, for example, on the quality, costs and outcomes of nursing (see Chapter 11). Furthermore, the optimality could be determined based on other consequences to nurses than merely their sick leave, for example, their occupational welfare and job satisfaction.

Part of the potential association between the RAFAELA system and HRM is indirect. The following is an example of the many indirect contributions of using a valid PCS. By using a creditable PCS, the organization may achieve fame. By making nursing workload visible, a PCS promotes and considers nurses primarily as a resource, not a cost. The hospital may be considered as modern and one that has become aware of the importance of human resources and thus takes care of its workers. That in turn may make recruitment of new employees easier for the hospital.

HRM stresses the importance of measuring essential HRM practices. Nursing intensity information should be included in the hospital's management support system. This data, however, is most informative and fruitful only after it has been integrated and related to other data systems supporting management, resembling, for example, the Balanced Scorecard (Kaplan & Norton 1992). The data from a questionnaire to nursing managers showed that not all of them have made use of the potential offered by the RAFAELA system as fully as possible (Pusa 2007). Similar conclusions of the insufficient administrative utilization of the results of the RAFAELA system were also made by Ojaniemi (2005).

The descriptions of the seven nursing intensity levels of the PAONCIL also include quality aspects. At the optimal level (the zero-level of the PAONCIL), where employees working capacity is used optimally, the quality of nursing care can best be maintained as according to the specification. At higher levels, the risk of mistakes increases and, thus, patient safety is endangered. Therefore, the beneficial significance of the RAFAELA system in the HRM of nurses is not limited only to the benefits of the employer and employees; patients do also benefit from it.

In conclusion, in the present situation of the insufficiency of both human and economic resources in health care and the contradiction between the supply and demand of health services, new solutions are badly claimed. How can one guarantee high-quality services, with favourable outcomes to the patients and to the satisfaction of all the parties: patients, employees, providers and financiers of health services? In the HRM of nurses, it is essential to ensure maximal performance while retaining an optimal workload. Likewise, it is important to maintain a good organizational culture, experience of justice in the distribution of work, job satisfaction and involvement of management and increase nurses' influence, empowerment and commitment. Part of the solution in reaching these aims is to measure nurses' workload with an adequately studied and scientifically validated patient classification system, for example, the RAFAELA system.

10 Conclusions

1. Based on previous research and on this study, the OPC instrument can be considered a well-validated and well-studied nursing intensity measure. It measures direct and indirect patient-associated workload in the wards of specialized health care sufficiently reliably for HRM, economic and scientific purposes. Instead of four to five nursing intensity categories, it is preferable to use direct nursing intensity points, six to 24 per patient per day. However, this scale with maximally a four-fold variation is scarcely completely sufficient to reflect the wide case mix and variation of patients' nursing intensity in the wards of specialized health care.

2. The PAONCIL measure replaces laborious time studies in the RAFAELA system as a link to the real world, that is, to nursing work time and, indirectly, also its costs. The analyses of its properties have shown it to be a valid measure. Compared to time studies, it is easy to use and it saves both time and money. As a bottom-up approach, it is based on the HRM concept of management.

3. The whole RAFAELA system has been proved to be a valid PCS, which can establish the current and the optimal nursing intensity creditably. The determination of optimal nursing intensity can be performed successfully in most wards, with a satisfactory explanatory power. Thus, as a whole, this classification system functions satisfactorily in the real world, in the context of Finnish specialized health care wards.

4. The limits of the RAFAELA system have been determined as to the explanatory power of the regression analysis, length of the PAONCIL follow-up period, response rate and upper limit of the average PAONCIL value, and so forth. Based on these analyses, it is now known under which circumstances this PCS is still valid.

5. The analysis of the significance of non-patient questions, concerning factors unrelated to the nursing activities of patients, showed that the OPC and the PAONCIL monitor quite the same phenomenon and are not markedly influenced by such factors unrelated to the nursing activities of patients.

6. The questions regarding non-patient factors can offer the HRM of nursing work valuable and feasible practical information on the functioning and problems of a ward.

7. After all these results, combined with existing previous research, the validity of the whole RAFAELA system and its elements as a measurement tool for HRM in nursing can be considered thoroughly analysed and evaluated.

8. Increasing the workload of nurses is linearly associated with their increased number of sick leaves. Among nurses with a workload >30% above the optimum, the rate of periods of sick leave is about 44 to 49% higher than among those with an optimum workload. This excess rate of sickness absence can result in about twelve extra sick leave days per person-year. Thus, part of the increased productivity resulting from work overload is lost in increased periods of sick leave. These findings suggest that work overload may significantly contribute to occupational health problems among employees and also lead to a considerable economic burden for employers.

9. The RAFAELA system has many features that, in principle, can support human resource management in nursing. The use of the RAFAELA system can be thought to support the central concept of HRM: high job satisfaction and high productivity simultaneously. Of course, the use of the bottom-up design RAFAELA system predominantly supports such aspects of HRM that are associated with workload measuring, planning of working conditions, increasing the involvement of managers and empowerment of nurses, increased feeling of justice. However, the RAFAELA is just a system of measures. Therefore, it can support the wide concept of HRM in nursing only to a certain degree and only in certain sectors. In practice, the potential for HRM offered by the RAFAELA system has not been utilized as fully as possible.

10. The following aspects can be presented as examples of the pure scientific contribution of this thesis. Firstly, in this thesis, the RAFAELA system has been proven as a valid PCS system. Thus, it can also be used as a measure for scientific research purposes to contribute to the production of additional scientific value. Secondly, the concept of optimal workload and the many possibilities for its determination have been illustrated, and the results motivate researchers to develop similar, scientifically validated measures for other occupational sectors, too. The principle of the PAONCIL instrument, the development of which is

described in this thesis, can also be used in these new measures, instead of time studies. Thirdly, the theoretical association of workload measurement tools and HRM has been analyzed and discussed. Fourthly, the association of measured excessive workload and increased sickness absence in nurses is a new scientific observation and additional work is needed to enlighten the association of workload with other aspects of occupational welfare. Thus, it can be considered that health care management science, nursing science and sciences investigating working life have obtained additional scientific value through this thesis.

11 Suggestions for further research

Although the RAFAELA system has an explanatory power between the OPC and the PAONCIL measures that can be considered as sufficient for the HRM purposes of nurses, it is probably possible to improve it further, as was discussed previously in Chapter 9.4. It may be possible to seek individual coefficients to all 24 weights of the OPC by using multiple regression analysis with a broad national material. Another possibility is to use the results of time studies at several hospitals as a basis for those 24 coefficients. However, it is important to maintain the uniform structure of the OPC measure between specialties and hospitals. Compromises must be sought between simplicity and uniformity on the one hand, and accuracy on the other.

Now that we have a measure of the optimal workload of nurses, it is important to study the association of excessive workload to other variables. We already know that excess average workload increases nurses' sick leave. But for how long a time must the excess workload continue before it causes consequences to nurses' state of health? Furthermore, is there latency between the period of excess workload and the increase in their sick leave? What is the significance of temporary peaks in workload, if they are followed by more peaceful periods? All these questions associated with sick leave are still unanswered.

Work overload causes increased sick leave, but it can also have other consequences to nurses' occupational welfare and job satisfaction. It is also possible to study these associations using the RAFAELA patient classification system, which allows the determination of optimal nursing intensity and workload. Work overload increases the risk of medical errors and adverse outcomes of patients. Using the RAFAELA system, it is now possible to study the association between a workload exceeding optimal and all those adverse events in more detail. It is also possible to study the association between nurses' work overload and patient satisfaction. How do patients experience the quality of medical and nursing care when nurses have an excessive workload? This is a wide-ranging issue that requires many more scientific studies.

References

Aaltonen J, Jaatinen S, Talvinko T, Virtanen M, Vohlonen I *et al*. 2007. Productization of the services of specialized health care in Finland. Suomen Kuntaliitto, Helsinki. (In Finnish).

Adams A, Bond S. 2003. Staffing in acute hospital wards: part 1. The relationship between number of nurses and ward organizational environment. Journal of Nursing Management 11, 287-292.

Aiken LH, Clarke SP, Sloane DM, Sochalski J, Silber JH. 2002. Hospital nurse staffing and patient mortality, nurse burnout, and job dissatisfaction. The Journal of the American Medical Association 288, 1987-1993.

Alward RR. 1983. Patient classification systems: the ideal vs. reality. Journal of Nursing Administration 13, 14-19.

Argyris C, Schön D. 1978. Theory in practice. Increasing professional effectiveness. Jossey-Bass, San Fransisco.

Arthur T, James N. 1994. Determining nurse staffing levels: a critical review of the literature. Journal of Advanced Nursing 19, 558-565.

Auld M. 1976. How Many Nurses? RCN, London.

Aydelotte M. 1973. Staffing for Quality Care. Journal of Nursing Administration February, 4-7.

Barnum BJ. 1990. Nursing theory analysis, application, evaluation. 3rd ed. Scott, Foresman and Co, Illinois.

Becker B, Huselid M. 2006. Strategic Human Resource Management: Where Do We Go From Here? Journal of Management 32(6): 898-925.

Botter M. 2000. The Use of Information Generated by a Patient Classification System. Journal of Nursing Administration 30(11):544-551.

Brewster C. 1993. Developing a “European” model of human resource management. The International Journal of human resource management 4(4), 765-784.

Brooten D, Youngblut J. 2006. Nurse Dose as a Concept. Journal of Nursing Scholarship 38(1), 94-99.

Brown CL. 1991. Caring in Nursing Administration: Healing through Empowering. In: Gaut DA, Leininger MM, (Eds). Caring: The Compassionate Healer. National League for Nursing Press , New York. 123-33.

Buerhaus PI., Staiger DO, Auerbach DI. 2000. Implications of an aging registered nurse workforce. The Journal of the American Medical Association 283, 2948-54.

Campbell T, Taylor S, Callaghan S, Shuldham C. 1997. Case mix type as a predictor of nursing workload. Journal of Nursing Management 5, 237-240.

Cardonna P, Tappen RM, Terrill M, Acosta M, Eusebe MI. 1997. Nursing staff time allocation in long-term care. A work sampling study. Journal of Nursing Administration 27, 28-36.

Carr-Hill RA, Jenkins-Clarke S. 1995. Measurement systems in principle and in practice: the example of nursing workload. Journal of Advanced Nursing 22, 221-225.

Chang E. 2002. Distributive justice and organizational commitment revisited: moderation by layoff in the case of Korean employees. Human Resource Management 41(2), 261-270.

De Groot HA. 1989. Patient classification system evaluation. Part 1: Essential system elements. Journal of Nursing Administration 19, 30-35.

De Groot HA. 1994a. Patient classification systems and staffing. Part 1, problems and promise. Journal of Nursing Administration 24, 43-51.

De Groot HA. 1994b. Patient classification systems and staffing. Part 2, practice and process. Journal of Nursing Administration 24, 17-23.

Ebener M.K. 1985. Reliability and validity basics for evaluating classification systems. Nursing Economics 3, 324-327.

Ellis J, Chapman S. 2006. Nurse staffing requirements. Nursing Management 13 (4), 30-33.

Fagerström L, Bergbom Engberg I. 1998. Measuring the unmeasurable: a caring science perspective on patient classification. Journal of Nursing Management 6, 165-172.

Fagerström L, Bergbom Engberg I, Eriksson K. 1998. A comparison between patients' experiences of how their caring needs have been met and the nurses' patient classification – an explorative study. Journal of Nursing Management 6, 369-377.

Fagerström L, Rainio A-K. 1998. Justifications for patient classification by OPC and an assessment of the reliability of the instrument. Proceedings of V Nationella vårdvetenskapliga konferensen: Vårdvetenskapens och vårdandets värld-kunskaps- och värdegrund. Vasa, Finland, 1998; 81-5. (in Swedish).

Fagerström L. 1999. The patient's caring needs. To understand and measure the unmeasurable. Doctoral thesis. Åbo Akademi University Press. Oy Fram Ab, Vaasa.

Fagerström L, Eriksson K, Engberg IB. 1999. The patient's perceived caring needs: measuring the unmeasurable. International Journal of Nursing Practice 5, 199-208.

Fagerström L, Rainio A–K. 1999. Professional assessment of optimal nursing care intensity – a new method for assessment of staffing levels for nursing care. Journal of Clinical Nursing 8, 369–379.

Fagerström L. 2000. Expert validation of the Oulu Patient Classification – a phase in the development of patient classification system, RAFAELA. Vård i Norden 20, 15-21

Fagerström L, Rauhala A. 2001. Finn Patient Classification – benchmarking in nursing care. A report of the pilot project from year 2000. Suomen Kuntaliitto, Helsinki, 2001. (in Finnish).

Fagerström L, Nojonen K Åkers A. 2002. Methodological triangulation as a validation method in the testing of the PAONCIL method. Journal of Nursing Science 14(4), 180-191.

Fagerström L, Rauhala A. 2003. Finn Patient Classification – benchmarking in nursing care. Final report of the project 2000 – 2002. Suomen Kuntaliitto, Helsinki. (In Finnish).

Fagerström L, Rauhala A. 2007. Benchmarking in nursing care by the RAFAELA patient classification system – a possibility for nurse managers. Journal of Nursing Management 15, 683-692.

Finnigan SA, Abel M, Dobler T, Hudon L, Terry B. 1993. Automated patient acuity. Linking nursing systems and quality measurement with patient outcomes. Journal of Nursing Administration 23, 62-71.

Gallagher JR. 1987. Developing a powerful and acceptable nurse staffing system. Nursing Management 18, 45-49.

Giovannetti P. 1979. Understanding patient classification systems. Journal of Nursing Administration 9, 4-9.

Giovannetti P, Mayer GG. 1984. Building confidence in patient classification systems. Nursing Management 15, 31-34.

Giovannetti P. 1986. Staffing methods - implications for quality. In: Williw LD, Linwood ME (Eds.). Measuring the Quality of Care. Churchill Livingstone, Singapore. 123-150.

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

Goldstone L, Ball J, Collier M. 1985. An Index of the Quality of Care for Acute Medical and Surgical Wards. Stockport Health Authority.

Guest D. 1989. Personnel and human resource management: Can you tell the difference. Personnel Management, January.

Hagerty BK, Spengler CD. 1985. Work sampling, analyzing nursing staff productivity. Journal of Nursing Administration 15, 9-14.

Harrison J. 2004. Addressing increasing patient acuity and nursing workload. Nursing Management 11 (4), 20-25.

Hegney D, Plank A, Parker V. 2003. Nursing workloads: the results of a study of Queensland Nurses. Journal of Nursing Management 11, 307-314.

Hendrickson G, Doddato TM, Kovner CT. 1990. How do nurses use their time? Journal of Nursing Administration. 20, 31-7.

Hinshaw A, Atwood J. 1982. A patient satisfaction instrument: precision by replication. Nursing Research 31, 170-175.

Hlusko DL, Nichols BS. 1996. Can you depend on your patient classification system? Journal of Nursing Administration 26, 39-44.

Hoffman F. 1988. Nursing Productivity Assessment and Costing Out Nursing Services. JB Lippincott Co, Philadelphia.

Huckabay LM, Skonieczny R. 1981. Patient classification systems: the problems faced. Nursing & Health Care 2, 89-102.

Hughes M. 1999. Nursing workload: an unquantifiable entity. Journal of Nursing Management 7, 317-322.

Hunt J. 1990. The activity balance. Nursing Standard 4(42), 47.

Iles P, Salaman G. 1995. Recruitment, selection and assessment. In: Storey J. (Ed.). Human resource management : a critical text. Routledge, London and New York. 203-233.

Janiszewski Goodin H. 2003. The nursing shortage in the United States of America: an integrative review of the literature. Journal of Advanced Nursing 43, 335-43.

Jennings BM, Rea RE, Antopol BB, Carty JL. 1989. Selecting, implementing, and evaluating patient classification systems: A measure of productivity. Nursing Administration Quarterly 17, 11-20.

Junnila, M. (Ed.). 2004. Sairaaloiden tuottavuus- Benchmarking -tietojen käyttö erikoissairaanhoidon toiminnan suunnittelussa, seurannassa ja arvioinnissa. Stakes, raportteja 280, Helsinki.

Kane RL, Shamliyan TA, Mueller C, Duval S, Wilt TJ, 2007. The association of registered nurse staffing levels and patient outcomes: systematic review and meta-analysis. Medical Care 45(12),1195-204.

Kaplan RS, Norton DP. 1992. The Balanced Scorecard – Measures that Drive Performance. Harvard Business Review January-February, 71-79.

Karma K. 1987. Käyttäytymistieteiden metodologian perusteet, 2nd ed. Otava, Keuruu.

Kaustinen T. 1995. The development and assessment of nursing care intensity classification at the Oulu University Central Hospital 1991-1993. Oulu University, Department of nursing science, Oulu. (In Finnish).

Keep E. 1989. Corporate training strategies: the vital component? In: Storey J. (Ed.). New Perspectives on Human Resource Management. Routledge, London and New York. 105-121.

Kelleher C. 1992. Validated indexes: key to nursing acuity standardization. Nursing Economics 10, 31-37.

Kinnunen J, Lindström K (Ed.). 2005. Rakenteellisen ja toiminnallisen muutoksen vaikutukset HUSin johtamiseen ja henkilöstön hyvinvointiin. Kuopion yliopiston julkaisuja E, Kuopion yliopisto.

Kinnunen J: (2007) Terveysalan johtamiskoulutus laajentunut ja monipuolistunut – Osaamisvaatimukset tuntuvat silti karkaavan. Premissi 2 (4), 12-16.

Kirk R. 1990. Using workload analysis and acuity systems to facilitate quality and productivity. Journal of Nursing Administration 20, 21-30.

Kivimäki M, Head J, Ferrie JE, Shipley MJ, Vahtera J, Marmot MG. 2003. Sickness absence as a global measure of health: evidence from mortality in the Whitehall II prospective cohort study. British Medical Journal 327, 364-369.

Kivimäki M, Forma P, Wikstrom J, Halmeenmäki T, Pentti J, Elovainio M Vahtera J. 2004. Sickness absence as a risk marker of future disability pension: the 10-town study. Journal of Epidemiology and Community Health 58, 710-711.

Kivimäki M, Head J, Ferrie J, Hemingway H, Shipley MJ, Vahtera J, Marmot MG. 2005. Working while ill as a risk factor for serious coronary events: the Whitehall II Study. American Journal of Public Health 95, 98-102.

Laine J. 2005. Quality and Productive Efficiency? An Economic Study on Institutional Care for Older People. Doctoral thesis. Gummerus Kirjapaino Oy, Saarijärvi.

Laine M. 2005. Hoitajana huomennakin. Hoitajien työpaikkaan ja ammattiin sitoutuminen. Doctoral thesis. Turun yliopiston julkaisuja C 233, Turku.

Lake WH. 1982. Nurse staffing based on Patient Classification. Information Management Service Inc., Rockville.

Lauharanta J, Virtanen M, Rotonen M, Tierala I, Kyllönen L, Kekomäki M.1997. NordDRG - erikoissairaanhoidon tuleva pohjoismainen tuotteistusstandardi? Suomen Lääkärilehti 35, 4210-4218.

Legge K. 1989. Human resource management: a critical analysis. In: Storey J. (Ed.). New Perspectives on Human Resource Management. Routledge, London and New York.20-40.

Levenstam A-K, Bergbom Engberg I. 1993. The Zebra system - a new patient classification system. Journal of Nursing Management 1, 229-237.

Levenstam A-K, Bergbom Engberg I. 2002. Journal of Nursing Management 10, 191-199.

Malloch K, Conovaloff A. 1999. Patient Classification Systems, Part 1. The Third Generation. Journal of Nursing Administration 29, 49-56.

Mark B. 2002. What explains Nurses' Perceptions of Staffing Adequacy? Journal of Nursing Administration 32 (5), 234-242.

Mayo AM, Van Slyck A. 1999. Developing Staffing Standards: Statistical considerations for Patient Care Administrators. Journal of Nursing Administration 29(10), 43-48.

McCallum A. 2004. The evaluation of hospitals' functioning in Great Britain. In Junnila, M. (Ed.). Sairaaloiden tuottavuus- Benchmarking -tietojen käyttö erikoissairaanhoidon toiminnan suunnittelussa, seurannassa ja arvioinnissa. Stakes, raportteja 280, Helsinki.(in Finnish).

McDowell I, Newell C. 1996. Measuring Health. A Guide to Rating Scales and Questionnaires. Oxford University Press, New York.

Melchior M, Niedhammer I, Berkman LF, Goldberg M. 2003. Do psychosocial work factors and social relations exert independent effects on sickness absence? A six year prospective study of the GAZEL cohort. Journal of Epidemiology and Community Health 57, 285-293.

Metsämuuronen J. 2003. Tutkimuksen tekemisen perusteet ihmistieteissä. 2nd ed. Gummerus Kirjapaino Oy, Jyväskylä.

Milne HA, McWilliam CL. 1996. Considering nursing resource as 'caring time'. Journal of Advanced Nursing 23, 810-9.

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), 463-71.

Nagaprasanna BR. 1988. Patient classification systems: strategies for the 1990s. Nursing Management 19, 105-112.

Needham J.1997. Accuracy in workload measurement: a fact or fallacy? Journal of Nursing Management 5, 83-87.

Needleman J., Buerhaus P., Mattke S., Stewart M. & Zelevinsky K. (2002) Nurse-staffing levels and the quality of care in hospitals. The New England Journal of Medicine 346, 1715-22.

Niedhammer I, Bugel I, Goldberg M, Leclerc A, Gueguen A. 1998. Psychosocial factors at work and sickness absence in the Gazel cohort: a prospective study. Occupational and Environmental Medicine 55, 735-741.

Nonaka I Takeuchi H (1995) The Knowledge Creating Company. Oxford University Press, Oxford and New York.

North FM, Syme SL, Feeney A, Shipley M, Marmot M. 1996. Psychosocial work environment and sickness absence among British civil servants: the Whitehall II study. American Journal of Public Health 86, 332-340.

Noyes B. 1994. Inter-rater reliability. Regaining credibility with your staff and financial officer while meeting JCAHO standards. Journal of Nursing Administration 24, 7-8.

Ojaniemi K. 2005. The administrative utilization of patient classification. Pro Gradu. Vaasa University, Vaasa. (In Finnish).

Ollila S. 2006. The Mastery of Strategic Competence-Based Management in Public and Private Social and Health Service Organizations: The Dimensions of Managerial Competence from the Viewpoint of Management Supervision. Doctoral Thesis. Vaasa University, Vaasa. (In Finnish).

Onnela E, Svenström R. 1998. Oulu patient classification system – further development in 1995-1997. Publication Series of the Northern Osthrobotnia Hospital District 4/1998, Oulu. (In Finnish).

Partanen P. 2002. Nurse staffing in acute care tertiary hospitals. Doctoral thesis. Kuopio University Publications, Kuopio. (In Finnish).

Perälä M-L. 1995. The quality assessment of patient care: validation of a quality measure (Qualpacs). Doctoral thesis. Gummerus Oy, Jyväskylä. (In Finnish).

Phillips CY, Castorr A, Prescott PA, Soeken K. 1992. Nursing Intensity Going beyond Patient Classification. Journal of Nursing Administration 22 (4), 46-52.

Procter S. 1992. Subjectivity and objectivity in the measurement of nursing workload. Journal of Clinical Nursing 1,123-9.

Procter S, Hunt M. 1994. Using the Delphi survey technique to develop a professional definition of nursing for analysing nursing workload. Journal of Advanced Nursing 19, 1003-14.

Pulkkinen M-L. 2000. Evaluation of nursing intensity on special care units. Doctoral thesis. University of Turku, Turku (In Finnish).

Pusa A-K. 2007. The Right Nurse in the Right Place. Nursing Productivity and Utilisation of the RAFAELA Patient Classification System in Nursing Management. Doctoral thesis. Kuopio University Publications, Kuopio.

Pyöriä P, Melin H, Blom R. 2005. Knowledge Workers in the Information Society. Evidence from Finland. Tampere University Press, Tampere.

Rainio A-K. 1994. Patient classification -- developing a method for measuring a patient's individual caring needs. Vasa sjukvårdsdistrikt skn, Vasa. (in Swedish).

Rauhala A, Rainio A-K, Fagerström L, Nojonen K. 1999. Hoitoisuusmittausten kehitystyön tuloksia ja käyttökokemuksia Vaasan keskussairaalasta. Suomen Lääkärilehti 29, 3571 – 3576.

Rees S, Bolton JPG, Mant JWF, Costa M. 1991. How many nurses?... a comparison of mechanisms for determining nurse staffing levels on hospital wards. Senior Nurse 11(5), 31-5.

Reitz JA. 1985. Toward a comprehensive nursing intensity index: part 1, development. Nurse Management 16, 21-30.

Rieder KA, Lensing SB. 1987. Nursing productivity: evolution of a systems model. Nursing Management 18, 33-44.

Rosenbaum HL, Willert TM, Kelly EA, Grey JF, McDonald BR. 1988. Costing Out Nursing Services Based on Acuity. Journal of Nursing Administration 18(7,8), 10-15.

Rothwell S. 1995. Human resource planning. In: Storey J. (Ed.) 1995. Human resource management : a critical text. Routledge, London and New York. 167-202.

Savolainen P, Töyry E. 1995. Hoitoisuusluokituskokeilu Monitor-hoitoisuusluokitusmittarilla Kuopion yliopistollisessa sairaalassa vuosina 1993-1995 (loppuraportti). Pohjois-Savon sairaanhoitopiirin julkaisuja No 28. Kuopion yliopistollinen sairaala, hallintokeskus. Kuopion yliopistollisen sairaalan monistamo, Kuopio.

Schroeder RE, Rhodes AM, Shields RE. 1984. Nurse acuity systems: CASH vs. GRASP: (a determination of nurse staff requirements). Nursing Forum 21(2), 72-7.

Seago J. 2002. A Comparison of Two Patient Classification Instruments in an Acute Care Hospital. Journal of Nursing Administration 32(5), 243-249.

Senge P. 1990. The fifth discipline. The art & practice of the learning organization. Doubleday/Currency, New York.

Seppälä A. 1992. Suomessa käytössä olevia hoitoisuusluokituksia. Opinnäytetutkielma. Terveydenhuollon koulutusohjelma. Kuopion yliopisto, Kuopio.

Sermeus W, Van den Heede K, Michiels D, Delesie L *et al*. 2004. A nation-wide project for the revision of the Belgian nursing minimum dataset: from concept to implementation. Studies in health technology and informatics 110, 21-6.

Sermeus V, Vleugels A, Van den Heede K, Diya L et al. 2007. Database quality of nursing care.

Shaha SH. 1995. Acuity systems and control charting. Quality Management in Health Care 3(3), 22-30.

SPRI . 1990. Vardtyngdsmätning. Exempel fran svensk sjukvard. Spri rapport 291, Stockholm.

Stenvall J, Majoinen K, Syväjärvi A, Vakkala H, Selin A. 2007. Mee romppeines siihen. Henkilöstövoimavarojen hallinta ja muutoksen johtaminen kuntafuusioissa. Acta, 191, Suomen kuntaliitto, Helsinki.

Stewart T. 1997. Intellectual Capital. The New Wealth of Organizations. Nicholas Brealey Publishing, London.

Storey J 1989. Introduction: from personnel management to human resource management. In: Storey J. (Ed.). New Perspectives on Human Resource Management. Routledge, London and New York.1-19.

Storey J. 1995. Human resource management: still marching on, or marching out? In: Storey J. (Ed.) 1995. Human resource management : a critical text. Routledge, London and New York. 3-32.

Streiner DL, Norman GR. 1995. Health Measurement Scales. 2nd ed. Oxford University Press, Oxford.

Strandmark Kjölsrud M. 1995. Vård i verkligheten om människovärde, maktrelationer och helhetssyn i professionell omvårdnad. Tiger, Malmö.

Strickland B, Neely S. 1995. Using a standard index to allocate nursing staff. Journal of Nursing Administration 25, 13-21.

Ståhle P & Grönroos M. 2000. Dynamic Intellectual Capital™. Knowledge Management in Theory and Practice. WSOY, Helsinki.

Sveiby KE. 1997. The new organizational wealth. Managing & measuring knowledge-based assets. Berrett-Koehler, San Francisco.

Sydänmaanlakka P. 2003. Intelligent leadership and leadership competencies. Developing a leadership framework for intelligent organizations. Doctoral thesis. Helsinki University of Technology. Monikko Oy, Espoo.

Sydänmaanlakka P. 2004. Älykäs organisaatio. Gummerus kirjapaino Oy, Jyväskylä.

Sädevirta J. 2004. Henkilöstöjohtamisen ja sen tutkimuksen kehittyminen. Työministeriön julkaisuja 35. Helsinki.

The Northern Nurses Federation (1995). Ethical guidelines for nursing research in the Nordic countries. Typotequet.

Tuomi J. 1997. The genesis of nursing and caring science in Finland. Doctoral thesis, University of Jyväskylä, Jyväskylä.

Turtiainen A-M. 1999. Methods to describe nursing with uniform language: The cross-cultural adaptation process of the Belgium Nursing Minimum Data Set in Finland. Doctoral thesis. University of Kuopio, Kuopio. (In Finnish).

Ulrich D. 1997. Measuring human resources: an overview of practice and a prescription for results. Human Resource Management 36(3), 303-320.

Vahtera J, Kivimäki M, Pentti J. 2001. The role of extended weekends in sickness absenteeism. Occupational and Environmental Medicine 58, 818-822.

Vahtera J, Pentti J, Kivimäki M. 2004. Sickness absence as a predictor of mortality among male and female employees. Journal of Epidemiology and Community Health 58, 321-326.

Vanputte AW, Sovie MD, Tarcinale MA, Stunden AE. 1985. Accounting for patient acuity: the nursing time dimension. Nursing Management 16, 27-36.

Van Slyk AA. 1991. Systems approach to the management of nursing services – part II: patient classification system. Nursing management 22(4),23-25.

Voutilainen P. 2004. The quality of nursing care in long-term care of older people. Doctoral thesis. Gummerus Oy, Saarijärvi. (In Finnish).

Viitala R. 2006a. Johda osaamista! Osaamisen johtaminen teoriasta käytäntöön. 2nd ed. Otavan kirjapaino Oy, Keuruu.

Viitala R. 2006b. Henkilöstöjohtaminen. 4th-5th ed. Edita, Helsinki.

Vingård E, Lindberg P, Josephson M, Voss M, Heijbel B, Alfredsson L, Stark S, Nygren Å. 2005. Long-term sick-listing among women in the public sector and its associations with age, social situation, lifestyle, and work factors: A three-year follow-up study. Scandinavian Journal of Public Health 33, 370-375.

Voss M, Floderus B, Diderichsen F. 2001. Physical, psychosocial, and organisational factors relative to sickness absence: a study based on Sweden Post. Occupational and Environmental Medicine 58, 178-184.

Voss M, Floderus B, Diderichsen F. 2004. How do job characteristics, family situation, domestic work, and lifestyle factors relate to sickness absence? A study based on Sweden Post. Journal of Occupational and Environmental Medicine 58, 178-184.

Vuori J (Ed.). 2005. Terveys ja johtaminen. Terveyshallintotiede terveydenhuollon työyhteisöissä. WSOY, Helsinki.

Welton J, Unruh L, Halloran E. 2006. Nurse Staffing, Nursing Intensity, Staff Mix, and Direct Nursing Care Costs Across Massachusetts Hospitals. JONA 36(9), 416-425.

Whitney JD, Killien MG. 1988. Establishing predictive validity of a patient classification system. Nursing Management 18, 80-86.

Williams M. 1988. When you don't develop your own: validation methods for patient classification systems. Nursing Management 19(3), 90-96.

Williams MA. 1977. Quantification of direct nursing care activities. Journal of Nursing Administration Oct: 15-18, 49-51.

Woods NF, Catanzaro M. 1988. Nursing Research Theory and Practice. C.V. Mosby, Toronto.

Zembala SA. 1993. Managing the costs of nursing care delivery. Nursing Administration Quarterly 17, 74-9.

Internet references

(Checked 17 August 2008)

1. European Commission Health & consumer protection directorate-general. Hospital activity in the European Union [www document]. <http://ec.europa.eu/health/ph_information/dissemination/hsis/hsis_11_en.htm>.

2. European Commission Health & consumer protection directorate-general; Directorate C - Public Health and Risk Assessment C2 - Health information. 16.01.2005. Qualified nurses and midwives per 100 000 inhabitants [www document]. <http://ec.europa.eu/health/ph_information/dissemination/hsis/human_4.pdf>.

3. Virtanen M, Pentti J, Vahtera J, Ferrie JE *et al.*. Overcrowding in Hospital Wards as a Predictor of Antidepressant Treatment Among Hospital Staff. *Am J Psychiatry*. Available at: <http://ajp.psychiatryonline.org/pap.dtl>.

Kuopio University Publications E. Social Sciences

E 141. Ensio, Anneli & Ryynänen, Olli-Pekka (toim.). Akuutin sydänpotilaan hoitoketjun arvioinnin kaksi näkökulmaa.
2007. 140 s.

E 142. Luukka, Katri. Vastavalmistuneiden lähihoitajien koulutuksenaikaiset merkitykselliset oppimiskokemukset vanhustyössä.
2007. 162 s. Acad. Diss.

E 143. Holopainen, Arja. Changes in meanings of teacherhood among finnish nurse teachers: a substantive theory of nurse teacherhood.
2007. 101 s. Acad. Diss.

E 144. Mäntynen, Raija. Kuntoutumista edistävä hoitotyö aivohalvauspotilaiden alkuvaiheen jälkeisessä moniammatillisessa kuntoutuksessa.
2007. 144 s. Acad. Diss.

E 145. Hautala-Jylhä, Pirjo-Liisa. Psychiatric Post-ward Outpatient Services – Between Hospital and Community.
2007. 112 s. Acad. Diss.

E 146. Kemppainen, Ulla. Ninth-grade adolescents´ health behavior in the Pitkäranta district (Russian Karelia) and in Eastern Finland.
2007. 149 s. Acad. Diss.

E 147. Pusa, Anna-Kaisa. The right nurse in the right place: nursing productivity and utilisation of the RAFAELA patient classification system in nursing management.
2007. 111 s. Acad. Diss.

E 148. Antikainen, Mari. Sosiaalityöntekijän asiantuntijuus lapsen huolto- ja tapaamissopimuspalvelussa.
2007. 199 s. Acad. Diss.

E 149. Hujala, Anneli. Johtamisen moniäänisyys: johtaminen vuorovaikutuksena ja puhuntana hoivayrityksissä.
2008. 146 s. Acad. Diss.

E 150. Kiviniemi, Liisa. Psykiatrisessa hoidossa olleen nuoren aikuisen kokemuksia elämästään ja elämää eteenpäin vievistä asioista.
2008. 119 s. Acad. Diss.

E 151. Mattila, Heleena. Voimaantumisen ydin. Sosiaali- ja terveysalalla toimivien ihmisten mahdollisuuksia voimaantua työssään.
2008. 119 s. Acad. Diss.

E 152. Martikainen, Janne. Application of decision-analytic modelling in health economic evaluations.
2008. 119 s. Acad. Diss.

E 153. Kaarakainen, Minna. Hajauttaminen valtion ja kuntien välisissä suhteissa 1945-2015: valtiollisesta järjestelmästä kohti kuntaverkostojen perusterveydenhuoltoa.
2008. 196 s. Acad. Diss.

E 154. Jäntti, Satu. Kansalainen terveyspalveluja valitsemassa: kolmivaiheinen valintamalli julkisissa ja yksityisissä lääkäripalveluissa.
2008. 199 s. Acad. Diss.