Federaal Kenniscentrum voor de Gezondheidszorg
Centre Fédéral d'Expertise des Soins de Santé
KCE

Financement des soins infirmiers hospitaliers

KCE reports vol. 53B

Federaal Kenniscentrum voor de Gezondheidszorg
Centre fédéral d'expertise des soins de santé
2007

Le Centre fédéral d'expertise des soins de santé

Présentation : Le Centre fédéral d'expertise des soins de santé est un parastatal, créé le 24 décembre 2002 par la loi-programme (articles 262 à 266), sous tutelle du Ministre de la Santé publique et des Affaires sociales, qui est chargé de réaliser des études éclairant la décision politique dans le domaine des soins de santé et de l'assurance maladie.

Conseil d'administration

Membres effectifs : Gillet Pierre (Président), Cuypers Dirk (Vice-Président), Avontroodt Yolande, De Cock Jo (Vice-Président), De Meyere Frank, De Ridder Henri, Gillet Jean-Bernard, Godin Jean-Noël, Goyens Floris, Kesteloot Katrien, Maes Jef, Mertens Pascal, Mertens Raf, Moens Marc, Perl François Smiets, Pierre, Van Massenhove Frank, Vandermeeren Philippe, Verertbruggen Patrick, Vermeyen Karel.

Membres suppléants : Annemans Lieven, Boonen Carine, Collin Benoît, Cuypers Rita, Dercq Jean-Paul, Désir Daniel, Lemye Roland, Palsterman Paul, Ponce Annick, Pirlot Viviane, Praet Jean-Claude, Remacle Anne, Schoonjans Chris, Schrooten Renaat, Vanderstappen Anne.

Commissaire du gouvernement : Roger Yves

Direction

Directeur général : Dirk Ramaekers

Directeur général adjoint : Jean-Pierre Closon

Contact

Centre fédéral d'expertise des soins de santé (KCE).
Rue de la Loi 62
B-1040 Bruxelles
Belgium

Tel: +32 [0]2 287 33 88
Fax: +32 [0]2 287 33 85

Email : info@centredexpertise.fgov.be
Web : http://www.centredexpertise.fgov.be

Financement des soins infirmiers hospitaliers

KCE reports vol. 53B

SERMEUS W., GILLET P., TAMBEUR W., GILLAIN D., GRIETENS J., LAPORT N., MICHIELS D., THONON O., VANDEN BOER G., VAN HERCK P., SWARTENBROEKX N., RAMAEKERS D

Federaal Kenniscentrum voor de gezondheidszorg
Centre fédéral d'expertise des soins de santé
2007

KCE reports vol.53B

Titre : Financement des soins infirmiers hospitaliers

Auteurs : Sermeus Walter (KULeuven), Gillet Pierre (ULg), Tambeur Wim UZLeuven), Gillain Daniel (ULg), Grietens Jan (UZLeuven), Laport Nancy ULg), Michiels Dominik (UZLeuven), Thonon Olivier (ULg), Vanden Boer Guy (UZLeuven), Van Herck Pieter (KULeuven), Swartenbroekx Nathalie, Ramaekers Dirk.

Relecture: Camberlin C., Peeters G., Mambourg F.

Experts externes : Coemelck Ddidier (RHMS, Tournai), Delvaux Anne (SPF Santé Publique), Folens Bert (FOD Volksgezondheid), Gobert Micheline (UCL), Lovinfosse André (CHR Citadelle, Liège), Meyers L.udo (Virga Jesse, Hasselt).

Acknowledgements The research team would like to thank Allen D., Bardiau F., Bartholomeyczik S., Berthou A., Defloor T., Folens B., Gobert M., Goossen W., Jylha V., Kelly M., Milisen K., Moen A., Moons P., Scott A., Straetmans D., Walshe M. and Welton J. for their significant contribution to the research project.

Validateurs : Darras Elisabeht, Junger Alain, Molenberghs Geert

Conflict of interest : Aucun conflit déclaré

Disclaimer: Les experts externes et validateurs ont collaboré à la rédaction du rapport scientifique mais ne sont pas responsables des recommandations aux Autorités. Les recommandations aux Autorités ont été rédigées par le Centre d'expertise (KCE).

Mise en Page : Nadia Bonnouh, Ine Verhulst

Bruxelles, 19 avril 2007

Etude nr 2005-19
Domaine : Health Services Research (HSR)

MeSH : Financial Management, Hospital; Economics, Hospital; Nursing Service, Hospital; Nursing; Financing; Evidence-based Medicine; Personnel Management

NLM classification : WY 125

Langage : français, anglais
Format : Adobe® PDF™ (A4)

Dépôt légal : D/2006/10.273/07

Comment citer ce rapport ?
Sermeus W., Gillet P., Tambeur W., Gillain D., Grietens J., Laport N, et al. Financement des soins infirmiers hospitaliers. Health Services Research (HSR). Bruxelles: Centre fédéral d'expertise des soins de santé (KCE); 2007. KCE reports 53B (D/2006/10.273/07).

Préface

La Belgique est l'un des rares pays où les activités infirmières contribuent à déterminer le financement de l'hôpital. Le niveau de ces activités est mesuré au moyen d'un instrument appelé Résumé Infirmier Minimum. Développé il y a plus de quinze ans, il est en cours de révision.

Le système actuel fait cependant l'objet de nombreuses critiques. La présente étude examine dans quelle mesure il serait possible d'allouer les moyens infirmiers de façon plus rigoureuse aux hôpitaux.

Déterminer de manière correcte les moyens nécessaires pour rencontrer les besoins en soins des patients dans un hôpital, est un exercice complexe. Est il possible de calculer de manière fiable le taux d'encadrement infirmier adéquat rien qu'à partir du résumé infirmier minimum? Le couplage de ces données avec les données de pathologie enregistrées par les médecins, peut il apporter une amélioration ?

Par ailleurs, pour la première fois à notre connaissance, un lien est établi avec l'evidence based nursing, ce qui offre des perspectives d'amélioration de la qualité des soins dans les hôpitaux belges.

Deux équipes universitaires, de l'ULg et de la KUL, ont collaboré avec le KCE pour attaquer cette question difficile. Les premiers résultats sont prometteurs et offrent des perspectives claires aux décideurs.

Jean-Pierre CLOSON
Directeur général adjoint

Dirk RAMAEKERS
Directeur général

Executive summary

INTRODUCTION

En Belgique, le système de financement des hôpitaux se singularise des autres systèmes de par le fait qu'il tient compte des soins infirmiers. Dans la plupart des autres pays, le coût moyen des soins infirmiers par jour est intégré dans les frais d'hôtellerie et de séjour. Cela signifie que les coûts infirmiers sont directement liés à la durée du séjour indépendamment des véritables besoins en soins des patients. Comme les coûts liés aux soins infirmiers représentent approximativement 50 % du budget total de personnel et 20 à 30 % des frais de fonctionnement d'un hôpital, cela peut mener à une grave compression des coûts, entraînant une surestimation des soins peu intensifs et une sous-estimation des soins très intensifs.

Bien que la Belgique tienne compte du coût des soins infirmiers dans ses systèmes de remboursement hospitaliers, le sentiment général est que cet ajustement n'est pas suffisamment précis. En effet, deux indicateurs de soins infirmiers, basés sur le Résumé Infirmier Minimum (RIM), sont utilisés dans le système de financement fédéral : un poids relatif (*cost-weight*) moyen pour les départements de chirurgie, de médecine interne et de pédiatrie et un ratio de soins intensifs pondéré (ZIP/ZAP) pour les départements de soins intensifs. Ce système de financement est critiqué : (1) il n'est pas lié aux DRG, (2) la pondération est basée sur les ratios d'effectifs réels en personnel, ce qui favorise les services ayant des niveaux élevés de personnel infirmier, (3) les poids relatifs ne semblent pas suffisamment sensibles aux changements dans les pratiques de soins infirmiers, (4) des unités avec une forte intensité de soins comme la gériatrie ne sont pas incluses dans le plan de financement complémentaire, (5) de nombreux incitants financiers visent exclusivement la réduction de la durée du séjour sans considération pour la compression des soins infirmiers durant ce même séjour.

L'utilisation de la première version du RIM dans le système actuel est également remise en question. Cette version, développée en 1985, constitue un enregistrement obligatoire depuis 1988 dans tous les hôpitaux aigus. Près de 20 ans plus tard, une question évidente est de savoir dans quelle mesure cette version du RIM est (encore) un instrument de mesure efficace des soins infirmiers hospitaliers pour différencier la dotation en personnel infirmier des besoins réels ? Une deuxième version du RIM a récemment vu le jour et s'intègre, aux côtés des résumés médicaux et autres, dans un ensemble plus vaste d'enregistrements de données, le RHM ou Résumé Hospitalier Minimal. Dans cette approche intégrée des banques de données disponibles, le RIM change également de nom et devient DI-RHM pour Données Infirmières du Résumé Hospitalier Minimum. Les dernières adaptations de la réglementation en vue de remplacer le RIM par le DI-RHM et de l'intégrer aux côté d'autres résumés au sein du RHM sont en préparation. L'implémentation officielle du DI-RHM est prévue en septembre 2007.

Le principal but de cette étude est d'examiner comment ce DI-RHM pourrait être utilisé et intégré dans le système de financement hospitalier tout en tenant compte des critiques susmentionnées. La principale préoccupation est qu'un système de financement hospitalier soit équitable et donne à chaque hôpital le budget dont il a besoin, compte tenu des caractéristiques des patients qu'il accueille et qu'il fournisse les ressources nécessaires pour offrir des soins sûrs. L'étude est exploratoire, destinée à mettre en lumière des voies présentant un potentiel de développement et d'utilisation dans le futur.

Cette étude vise à examiner les questions de recherche suivantes :

- Quels outils et méthodes sont utilisés à l'étranger pour le financement des soins infirmiers hospitaliers ? Quelles sont les caractéristiques des méthodes utilisées ?
- Quelles données probantes peuvent être trouvées dans la littérature concernant une série d'interventions infirmières enregistrées dans le DI-RHM ?

- Comment le DI-RHM doit-il être utilisé dans un système de financement des soins infirmiers hospitaliers ?

MÉTHODES

Analyse de la littérature

Cette étude est divisée en quatre grandes parties. La première partie consiste en une revue de la littérature à propos de la manière dont les soins infirmiers sont intégrés dans les différents systèmes de financement des hôpitaux. Exploitant la littérature « grise » qui se dissimule dans les rapports gouvernementaux, notre étude utilise les résultats du projet « HEALTHBASKET » financé par l'Union Européenne qui a été réalisé par l'European Health Care Management Association (EHCMA) dans neuf pays européens et une enquête sur le financement des hôpitaux réalisée par l'European Hospital and Healthcare Federation HOPE dans les Etats membres de l'Union Européenne. Elle a été complétée par une enquête réalisée parmi les membres des associations suivantes : Patient Classification Systems International (PCSI), International Medical Informatics Association (IMIA), Nursing Informatics Workgroup (IMA-NI) et les représentants nationaux de l'European Federation of Nurses (EFN). Sur 17 contacts, sept ont répondu.

Evaluation du niveau de preuve

La deuxième partie de cette étude a consisté à évaluer le niveau de preuve des interventions répertoriées dans le DI-RHM. Neuf interventions ont été choisies, en fonction de leur fréquence (1) et de leur variabilité d'occurrence (2) dans les hôpitaux belges, de leur relation possible avec la dotation en personnel infirmier (3) et de la mise en évidence de données probantes pour l'intervention infirmière sélectionnée (4). Les critères un et deux ont été testés sur les données collectées durant la phase pilote du DI-RHM. Durant ce projet, 117395 observations ont été réunies dans 66 hôpitaux et 231 unités de soins infirmiers. Les critères trois et quatre ont été cotés de 1 à 5 par un panel de 7 experts en soins infirmiers. Sur la base de ces quatre critères, neuf interventions infirmières ont été sélectionnées. Pour ces interventions, différentes stratégies de recherche ont été développées. Les principales sources ont été des recommandations de pratique basée sur l'evidence, des revues systématiques et des publications spécifiques EBN (Evidence-Based Nursing) provenant de sources « fiables » comme NICE, SIGN, CBO, JBI, WVVH, NCCHTA, Duodecim, CEBAM-LIBRARY, CDSR, DARE, Clinical Evidence, Evidence based nursing et ICSI. Lorsque suffisamment de données probantes ne pouvaient pas être trouvées parmi ces différentes sources, des études originales ont également été incluses. Toutes les publications ont été triées systématiquement en fonction de la compatibilité avec la définition de la question de recherche spécifique de l'intervention DI-RHM, à l'aide d'une approche PICO[1]. Une évaluation de la qualité méthodologique des publications a été réalisée et, après une extraction systématique de données, un ensemble de recommandations Evidence-Based (E-B) avec indication du niveau de preuve A, B ou C a été constitué pour chaque intervention infirmière retenue.

Pour l'une des interventions (prévention des escarres), une arborescence EBN a été développée permettant d'interroger la base de données RCM pour tester le niveau de preuve des interventions infirmières. Pour chacun des éléments de cette structure en arbre, la disponibilité de données cliniques pertinentes ainsi que le codage ICD–9 original dans les résumés RCM/RIM et DI-RHM ont été pris en compte. L'algorithme a été programmé comme un ensemble minimal de règles, utilisant SAS version 9.1®. L'ensemble de ces règles a été testé sur les résumés RCM/RIM et DI-RHM couplés disponibles évaluant la prévention des escarres chez 6030 patients.

[1] Des mots clés sur les niveaux 'Patient' 'Intervention' 'Comparison' 'Outcome' sont utilisés pour faire des recherches dans les bases de données

Cas de patients

La troisième partie de cette étude a consisté à rédiger 112 cas cliniques réels de patients. Ces cas ont été collectés dans 35 hôpitaux. Chaque cas a été rédigé dans un vocabulaire clinique de sorte que les besoins de personnel y relatifs puissent être évalués par des infirmiers et des infirmiers-chefs. Quelques questions simples ont été posées : si vous deviez effectuer les soins de ces patients, combien de temps cela vous prendrait-il ? De combien de ces patients pourriez-vous vous occuper ? Si vous n'aviez pas de limitations de ressources, quelle différence cela ferait ? Les différentes questions permettent d'évaluer la cohérence interne de la procédure de cotation. Ces cas ont été distribués de manière aléatoire parmi les infirmiers de sorte que chaque infirmier a dû coter en moyenne 10 cas et chaque cas a été évalué en moyenne par 8 infirmiers. Ces infirmiers ne connaissaient pas le patient en question et ne travaillaient pas dans l'hôpital où le cas a été rédigé. 202 infirmiers de 69 hôpitaux ont participé à cette étude pour coter ces cas. Au moment de la rédaction des cas, les résumés infirmiers minimum (RIM et DI-RHM) ainsi que certains systèmes éprouvés de classification de patients, comme le TISS (Therapeutic Intervention Scoring System) en Soins Intensifs, la grille AGGIR (Autonomie Gérontologique Groupes Iso Ressources) en gériatrie, la classification San Joaquin et l'indice pédiatrique NARVEL (Nursing Attention Requirement Level), ont été scorés. La comparaison des scores DI-RHM avec ces autres systèmes de classification permet d'évaluer la cohérence externe de la procédure de cotation. Indépendamment de la cotation des cas, chacune des 79 interventions infirmières du DI-RHM a été cotée par 20 personnes sélectionnées de manière aléatoire, évaluant le temps nécessaire pour réaliser chacune de ces interventions. En raison de la disponibilité du profil DI-RHM par patient, un score « sum_intervention » par cas pouvait être calculé. Ce score a été comparé à la cotation du temps par cas à titre de mesure de validité prédictive. Pour les deux cotations (cas clinique et intervention DI-RHM), un estimateur robuste (moyenne de Huber) a été utilisé comme mesure pour la tendance centrale lorsque les données étaient très déviées. Cinq cas ont été réécrits à l'aide des recommandations EBN des neuf interventions infirmières qui avaient été étudiées en vue de rassembler des données probantes. Les soins réels ont été remplacés par les soins nécessaires sur base de toutes les données probantes (A à C). Ces 5 cas ont été soumis à 10 autres personnes sélectionnées de manière aléatoire (différentes de celles qui avaient coté les cas originaux). Les scores de dotation en personnel des cas modifiés sur base de données probantes ont été comparés avec les scores de dotation en personnel des cas originaux.

DRG et coûts des soins infirmiers

La quatrième partie de cette étude a consisté à relier les DRG et les coûts des soins infirmiers. Les coûts des soins infirmiers ont été mesurés dans six catégories de durée de soins infirmiers. La durée des soins infirmiers par patient par jour a été mesurée à l'aide du DI-RHM en y ajoutant les points de durée relatifs par intervention cotée. Pour cette partie de l'étude, le résumé collecté durant la phase pilote du DI-RHM a de nouveau été utilisé. Parmi les données RCM et RIM disponibles, 60019 observations ont pu être couplées. Cet échantillon a été subdivisé en deux autres sous-échantillons aléatoires, l'un pour construire le modèle, l'autre pour tester sa validité et sa stabilité. Différents modèles ont été testés. Finalement, un modèle de régression logistique multinomial a été utilisé pour l'analyse des données.

RÉSULTATS

Il est généralement admis dans la littérature que les différences de soins infirmiers sont faiblement expliquées par les DRG. Le coefficient de détermination varie entre 20 % et 40 %. Des coefficients de variance élevés par DRG ont été rapportés. La part des soins infirmiers dans les frais totaux par DRG varie entre 6 % et 25 %. Cette grande variation a déjà été identifiée dès le début par les principaux développeurs des DRG, B. Fetter et J. Thompson. Plusieurs projets de recherche ont d'ailleurs déjà traité ces questions. Malgré tout, le résultat de ceux-ci reste faible, principalement parce que les relations ne

sont pas encore bien comprises. Comme J. Thompson & D. Diers (1992) l'ont dit : « Les comparaisons entre les hôpitaux suggèrent qu'il y a encore tellement de choses à comprendre à propos des différences d'intensité en soins infirmiers que tout changement de la politique de remboursement, comme la pondération des DRG par l'intensité des soins infirmiers, serait prématuré ».

Dans l'analyse de la littérature sur la manière dont les coûts de soins infirmiers sont pris en compte dans les systèmes de remboursement, cinq systèmes différents ont été identifiés (voir Schéma 1) :

1a) Pays utilisant les DRG sans les adapter pour tenir compte des soins infirmiers comme les Pays-Bas, le Royaume-Uni et l'Italie ;

1b) Pays utilisant les DRG sans les adapter pour tenir compte des soins infirmiers, mais avec des projets de prise en compte des soins infirmiers dans le financement comme le Danemark et les États-Unis ;

1c) Pays utilisant les DRG sans les adapter pour tenir compte des soins infirmiers aujourd'hui, mais qui l'ont fait dans le passé, comme la France et l'Allemagne ;

2) Pays utilisant les DRG en les adaptant pour tenir compte des soins infirmiers, comme le Canada, l'Australie et la Nouvelle-Zélande ;

3) Pays utilisant les DRG en les adaptant pour tenir compte des soins infirmiers et en calculant le coût réel des soins infirmiers par patient ce qui devrait permettre de mesurer la variabilité des frais de soins infirmiers à l'intérieur de chaque DRG, comme la Suisse ;

4) Pays adaptant le système de financement hospitalier pour tenir compte des soins infirmiers, mais pas directement en fonction des DRG, comme la Belgique et le Luxembourg ;

5) Pays n'utilisant pas de DRG et ne tenant pas compte non plus des soins infirmiers.

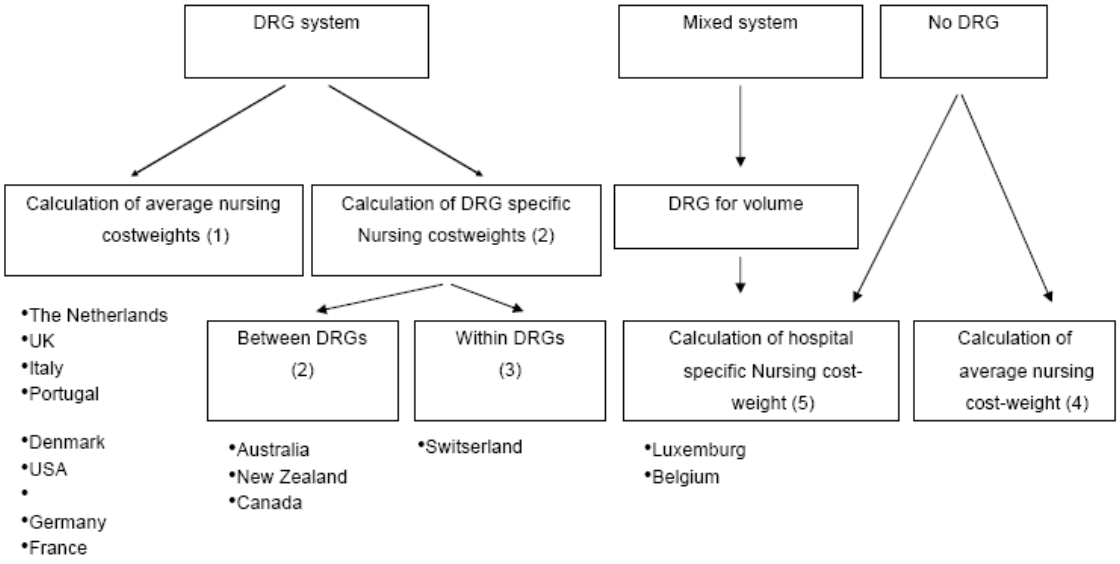


Schéma 1 : Coût de soins infirmiers dans le système de remboursement

La manière dont le Canada ou l'Australie tiennent compte des coûts des soins infirmiers est assez exemplaire. Ils utilisent les DRG pour décrire les cas traités (*casemix*). En outre, un système d'allocation des coûts est utilisé, dans lequel une liste de centres de frais est définie. Pour chaque centre de frais, des facteurs de coûts sont identifiés. Le facteur de coûts pour les soins infirmiers est la durée des soins infirmiers définie par un système de classification des patients. Sur la base de ce facteur de coûts, le coût moyen des soins infirmiers par DRG est identifié ; si aucun système de classification des

patients n'est utilisé, la lourdeur des soins infirmiers a été identifiée pour relier la durée des soins infirmiers aux DRG.

Comme le résultat est toujours une durée moyenne de soins infirmiers par DRG, la Suisse va un pas plus loin en associant la charge de travail des infirmiers et les données DRG de manière continue, en tirant ces données du dossier électronique du patient. Aux États-Unis, il existe des plans pour la facturation de l'intensité des soins infirmiers qui doivent permettre de prendre en compte la variabilité dans un DRG.

La principale condition est la disponibilité des données de soins infirmiers. C'est à nouveau la principale critique aux États-Unis pour tenir compte des soins infirmiers dans la révision du système de remboursement hospitalier par APR-DRG. Même si un résumé infirmier minimal a été défini par Werley en 1985, aucun ensemble systématique et comparable de données de soins infirmiers par admission n'est disponible.

Dans une seconde partie de cette étude, le niveau EBN de neuf interventions infirmières du DI-RHM a été recherché. La plupart des recommandations sont d'un niveau de preuve C. Quelques preuves limitées sont de niveau A ou B. Cela signifie que le développement d'un système de remboursement hospitalier fondé sur des pratiques infirmières E-B et non sur les pratiques infirmières réelles sera difficile en raison du manque de preuves disponibles. Un outil important est le développement d'un ensemble de règles EBN pour la prévention des escarres. L'algorithme a été appliqué sur un ensemble limité de 6030 patients des bases de données infirmières et médicales couplées. Les résultats montrent 1,3 % d'« excès de soins » avec un matelas adapté et 2,6 % d'excès de soins en ce qui concerne les changements de position. En outre, sur les 6030 patients, 1335 (22,1 % du total) auraient dû bénéficier d'un matelas dynamique. 1054 patients (17,5 % du total) n'ont pas reçu les soins requis. Cela constitue une mesure du « manque de soins ». De même, dans 28,4 % des cas, un manque d'éducation relative à la prévention des escarres a été identifié.

La troisième partie de cette étude était la plus importante. Cent douze cas cliniques réels ont été décrits et utilisés pour valider l'estimation de charge de travail des infirmiers par le DI-RHM. L'un des principaux résultats de cette étude est que des pondérations valides, fiables et utilisables des soins infirmiers par intervention DI-RHM ont été développées. Des estimateurs « robustes » ont été utilisés pour tenir compte des grandes différences entre les personnes attribuant des scores.

Les pondérations de soins infirmiers ont été validées pour les 112 cas cliniques. Il y a une grande corrélation (r=0,90) entre la somme des pondérations des soins infirmiers par intervention et la cotation directe de la durée des soins infirmiers. La corrélation entre les nouvelles pondérations des soins infirmiers développées pour DI-RHM et les pondérations du Professeur Closon et de l'Université de Gent pour le RIM est de plus de 0,93. La durée calculée pour les cas montre une grande corrélation avec les systèmes de classification des patients validés comme le TISS, San Joaquin et AGGIR.

La corrélation avec l'indice NARVEL n'est pas significative. Celui-ci, développé uniquement pour les services pédiatriques en 1975, n'est par ailleurs plus utilisé ni validé.

Il y avait une petite différence significative, dans les cotations des cas entre les régions francophone et néerlandophone. Ces différences sont prévisibles et peuvent être liées aux différences d'environnement, de conditions de travail, d'organisation, de mélanges de compétences, de perceptions en matière d'encadrement infirmier, etc…

Il n'y avait pas de différence significative dans la codification des cas entre les cas cliniques réels et les cas corrigés pour l'EBN. Cela signifie que, du point de vue de la dotation en personnel infirmier, les soins EBN ne sont pas toujours plus ou moins coûteux. Cela signifie également que les décisions de dotation en personnel sont probablement peu précises et ne peuvent pas être évaluées en quelques minutes de plus ou de moins, mais plutôt en fonction des soins à donner à un patient en plus ou en moins.

Dans la quatrième phase, un modèle a été développé pour lier les DRG aux données de soins infirmiers. Le principal problème était que, compte tenu de la conception de

l'échantillonnage, seules 15 % des données de soins infirmiers étaient disponibles et 85 % ont dû être estimées sur la base du RCM. Sur base de l'analyse de la littérature, plusieurs variables ont été incluses : DRG et gravité de la maladie, âge et sexe du patient, soins intensifs versus non intensifs, soins chirurgicaux versus non chirurgicaux, spécialité médicale, soins de routine versus soins d'urgence, durée du séjour et jour précis dans le séjour. Les DRG ont été groupés en six catégories qui se sont avérées les plus homogènes pour les soins infirmiers. Le principal résultat est une variance expliquée d'environ 40 %. 38 % de toutes les journées ont été classées correctement par le modèle.

Mais il s'est avéré que l'échantillon disponible n'était pas représentatif de tous les DRG. Seule une unité de soins infirmiers sur cinq par hôpital a été incluse dans l'échantillon, de sorte que le modèle ne pourrait pas être comparé avec l'actuel système de financement à un niveau national.

CONCLUSIONS

Une première conclusion de cette étude est qu'il est possible de pondérer les soins infirmiers sur base du niveau requis de dotation en personnel plutôt que sur base des niveaux observés. Cette étude montre que ce n'est pas seulement possible, mais que cela offre également des pondérations des soins infirmiers valides, fiables et utilisables. Elles ont été validées par rapport à 112 cas cliniques réels. La disponibilité de ces cas est un atout majeur de cette étude, car cela permet également d'évaluer les niveaux de personnel belge requis par rapport aux niveaux internationaux (p.ex. Pays-Bas, France, Suisse).

Une deuxième conclusion est que les données probantes en matière de soins infirmiers sont limitées. Du point de vue de la dotation en personnel infirmier, il n'y a pas de véritable différence dans l'évaluation des besoins de personnel entre les soins E-B et les soins infirmiers observés. Il convient de relever que, à partir de l'analyse de la littérature sur les escarres, une structure arborescente a pu être développée et pourrait être utilisée comme ensemble minimum de règles pour vérifier la pertinence des interventions des hôpitaux et des infirmiers en matière de prévention des escarres. Le développement de cet ensemble de règles combinant les données infirmières et médicales constitue une perspective intéressante pour une recherche ultérieure sur la manière dont des soins plus E-B pourraient être inclus dans le système de remboursement des hôpitaux. Le lien avec des systèmes *pay for performance* (P4P) ou *pay for quality* (P4Q) est ici tout à fait évident.

Une troisième conclusion est que les données de soins infirmiers peuvent être liées aux DRG, mais il faut encore beaucoup travailler avant de pouvoir valider le modèle. L'impact du modèle sur un système de financement hospitalier complet à un niveau national n'a pas pu être testé. Il faudrait tester le modèle sur des données couplées RCM-RIM pendant 3 années consécutives. Le financement actuel devrait alors être comparé avec le nouveau modèle développé.

Une quatrième conclusion est que le résultat de cette étude, soit 6 catégories de coûts de soins infirmiers par DRG, est plus transparent pour les utilisateurs et les décideurs que les méthodes de financement actuelles utilisant les zones, ZIP/ZAP, les déciles, etc. Le fondement statistique permettant de déduire ces six catégories de la manière la plus appropriée est assez complexe, mais le résultat est facile à lire et à comprendre. Chaque hôpital peut comparer son propre profil de soins infirmiers par DRG avec le profil national. Si d'autres ensembles de règles pouvaient être développés et testés sur un échantillon plus large, des profils E-B pourraient être générés et aideraient les hôpitaux à se comparer à un benchmark plus « EBN ». Le système actuel de remboursement des soins infirmiers hospitaliers n'incite pas à changer les pratiques. Le lien entre DRG et EBN aiderait à fournir plus d'incitants à la qualité et à l'efficacité.

Une cinquième conclusion est que la liaison entre DRG et données de soins infirmiers contribuerait à mettre en œuvre un ajustement des soins infirmiers au remboursement hospitalier à l'échelle de l'hôpital. Dans le plan de financement actuel, l'ajustement pour

tenir compte des soins infirmiers est limité aux unités de chirurgie, de médecine interne, de soins intensifs et de pédiatrie. Il n'y pas d'ajustement pour les unités de gériatrie même si les soins infirmiers sont l'une des principales caractéristiques des soins du patient en gériatrie. Une liaison aux DRG ferait moins dépendre le remboursement des structures et des départements, et proposerait une évolution vers un financement par patients et par programme de soins.

Une sixième conclusion qu'il existe de nombreuses alternatives quant à la manière d'intégrer la composante « soins infirmiers » dans le système de remboursement des hôpitaux. Une première approche est que les données des soins infirmiers soient utilisées pour un calibrage annuel des pondérations des coûts de soins infirmiers des DRG. Cela signifie que 15 % des données de soins infirmiers réelles existantes seraient utilisées pour estimer le modèle de 100 % de tous les séjours. Le principal avantage de cette approche est qu'il n'y a pas d'impact direct des scores sur le financement de sorte que l'« effet pervers » de l'enregistrement des soins infirmiers serait limité. En effet, une sur cotation ou une sous-cotation entraînera uniquement un ajustement du modèle qui s'appliquera à tous les hôpitaux aigus. D'autre part, cela n'entravera pas l'effet pervers de l'enregistrement des RCM qui pourrait mener à un remboursement plus important. Le principal désavantage est qu'une pondération des coûts nationaux moyens par DRG est calculée et que la variabilité à l'intérieur des DRG est ignorée. Il se pourrait qu'un hôpital ait une pondération de coûts plus élevée pour un DRG donné qu'un autre hôpital parce que sa durée de séjour est réduite avec des soins plus intensifs ou qu'il a tendance à avoir des patients avec des besoins en soins infirmiers plus importants. Cela ne serait pas pris en compte, entraînant une distorsion du financement des besoins de personnel infirmier. Une seconde approche pourrait être que le profil de soins infirmiers actuels par DRG par hôpital soit pris en compte. Cette approche est plus sensible à l'effet pervers, mais probablement plus proche des différences de pratiques de soins infirmiers. Le principal inconvénient est que nous ne sommes pas sûrs de la manière dont les 15 % de l'échantillon représentent les 100 % des soins infirmiers par DRG. Les DRG à volume élevé seront probablement bien représentés. C'est moins clair en ce qui concerne la représentation des DRG à faible volume et les cas extrêmes.

Le système de remboursement final pourrait être un mélange entre les soins infirmiers réels et modélisés.

Cette étude avait plusieurs limitations.

Une première limitation est que cette étude était exploratoire. Seules neuf des 78 interventions DI-RHM ont été examinées en termes d'EBN. Seul un ensemble de règles (pour la prévention des escarres) a été développé. Seuls des cas de soins chirurgicaux, médicaux, pédiatriques, gériatriques et intensifs ont été rédigés et cotés. Cette étude a montré que de nouveaux investissements en EBN pour le développement d'ensembles de règles pourraient être utiles et qu'une extension à d'autres cas cliniques (soins de maternité, soins néonataux, soins chroniques) est à recommander.

Deuxièmement, il y a eu des limitations majeures dans les données disponibles. Tout d'abord, la base de données utilisée pour cette étude a été celle qui a été constituée au cours du développement et des phases test du DI-RHM. Il est évident qu'il n'y avait pas d'alternative car, à ce stade, aucune autre base de données DI-RHM n'était disponible. De plus, la version finale du DI-RHM n'est pas équivalente à celle utilisée dans ces phases de développement et de test. Par ailleurs, les infirmiers, bien que sensibilisés, n'étaient pas habitués aux nouvelles définitions et données. Enfin, les données n'étaient pas représentatives de tout l'hôpital car maximum cinq unités de soins par hôpital ont été impliquées dans cette étude. Au final, l'ensemble de données a été utile dans une perspective exploratoire pour tester différentes manières de coupler les données. Cependant, une nouvelle validation du modèle est requise. Différentes alternatives et leur impact sur le remboursement des hôpitaux doivent être examinés sur un échantillon représentatif de données hospitalières.

Troisièmement, l'analyse des données est un problème statistique complexe, compte tenu des données imbriquées et autocorrélées, des durées de séjour variables et des 85 % de données manquantes. Un modèle logistique multinomial robuste a été utilisé pour le modèle. Il faut examiner si l'échantillon actuel de 15 % est représentatif au niveau DRG (ou de groupes de DRG) et si d'autres méthodes statistiques plus pointus amélioreraient le caractère prédictif du modèle.

RECOMMANDATIONS

Ce projet relatif au financement des soins infirmiers en milieu hospitalier doit être considéré comme une étude de faisabilité. Toute une série d'alternatives ont été évaluées, parmi lesquelles certaines semblent mériter un complément d'investigation. Le rapport pose les premières pierres d'une application éventuelle de système de financement hospitalier.

Bien que la Belgique ajuste son système de financement hospitalier en fonction du coût des soins infirmiers, l'impression générale est que cet ajustement n'est pas suffisamment précis. Le KCE recommande d'introduire un mode de prise en charge du staff infirmier plus adéquat, sur base de taux d'encadrement appropriés et d'une pondération par les coûts salariaux infirmiers, ce qui équivaut à une critique du système actuel.

L'étude montre qu'il est possible de développer une mesure des besoins d'encadrement à partir du RIM qui soit utilisable pour le financement des soins infirmiers. De plus, les données de soins infirmiers ont été couplées aux DRGs, en utilisant six types de coûts infirmiers par DRG. L'impact du modèle sur l'ensemble du système de financement hospitalier au niveau national, n'a pas pu être testé. Il serait utile de le faire sur les données couplées du RIM et du RCM pendant une période de trois années consécutives et de comparer les résultats du modèle actuel de financement avec ceux du modèle développé. Le présent rapport donne aux décideurs un outil pour apprécier l'intérêt d'un modèle alternatif. Le KCE recommande de poursuivre des études de validation de ce modèle prometteur, si les décideurs ont l'intention de s'engager dans cette voie.

Le KCE recommande que toutes les écoles d'infirmières incorporent dans leurs programmes « l'evidence based nursing » et y insistent sur la notion de « soins appropriés ». Cela commence déjà à être le cas mais pas de manière suffisamment intégrée à l'ensemble du programme.

Le KCE recommande que les ensembles de règles evidence-based soient développés davantage et appliqués aux données, de façon à pouvoir produire des profils evidence-based et à aider les hôpitaux à se comparer au moyen d'un benchmark plus « EBN ». Un lien entre le RIM, les DRG et l'evidence-based nursing pourrait, en donnant des benchmarks de meilleure qualité et plus efficients, constituer un incitant à des soins de meilleure qualité.

Table of contents

Introduction

The current Belgian financing system of hospital nursing care is not regarded as sufficiently accurate in the allocation of resources between Belgian general hospitals. Two nursing care indicators that are based on the Belgian Nursing minimum Dataset (B-NMDS) are used in the Belgian financing system: an average cost-weight for surgical, internal medicine and paediatrics departments and a weighted intensive care ratio (ZIP/ZAP) for intensive care departments. This financing system is criticized: (1) it is not linked with DRGs, (2) cost-weighting is based on actual staffing ratios, which favours nursing wards with high nurse staffing levels, (3) cost-weights seem not sensitive enough for changes in nursing practice, (4) nursing intensive departments such as geriatrics are not included in the complementary financing scheme, (5) many financial incentives are focused on reducing the length of stay without considering the compression of nursing care during that stay.

The use of B-NMDS version I (B-NMDS-I) within this current system is also questioned by multiple stakeholders such as hospital administration and professional healthcare workers. This version was developed in 1985 and implemented in 1988. It is an obvious question in 2007, to which degree this NMDS version is (still) a good measure to differentiate hospital nursing care between settings in its staffing and resource needs? The B-NMDS has recently been updated to B-NMDS version II (B-NMDS-II) in a large research project, granted by the Belgian Federal Public Service for Health, Food Chain Safety and Environment [1]. The creation of this renewed instrument focused on multiple applications: hospital staffing policy, continuous quality improvement by means of clinical indicators, justification of hospital admission and length of stay using Appropriateness Evaluation Protocols (A.E.P.), and finally: application within the financing system of hospital nursing care. The final legal regulations to replace B-NMDS-I by the B-NMDS-II and to integrate the dataset together with the medical and other datasets in the hospital discharge dataset (Minimale Ziekenhuisgegevens: MZG) are in preparation. The final implementation of the new dataset is foreseen in September 2007. The main goal of this study is to investigate how this B-NMDS-II could be used and integrated in the hospital financing system and reconcile the critiques posed above. The main concern is that a hospital financing system should be fair and give every hospital the budget it needs to give the care that corresponds with the needs of patients and would provide the resources that are needed to give safe care.

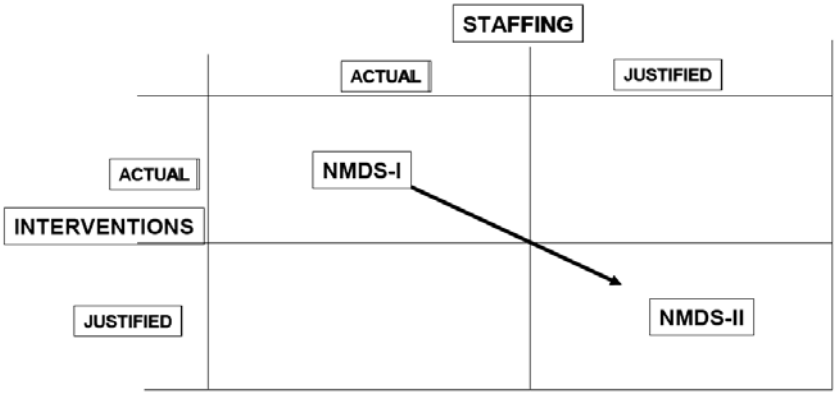
This study aims to investigate following research questions:

- Which tools and methods are used abroad in financing hospital nursing care? Which are the characteristics of the methods used?
- Which evidence can be found in literature concerning a series of nursing interventions as registered in NMDSII?
- How should NMDSII be used in a hospital nursing financing system?

The study is explorative, aimed at discovering ways with potential for further development and use. It also takes practical feasibility issues into account. The first research question draws on lessons learned out of international experience. Countries all over the world are tackling the issue of hospital nurse costing. Their experiences and alternative approaches should be considered to add extra value to the financing model.

The main model that will be followed in the study is given by figure 1:

Figure 1: A two way approach towards better financing of nursing care



As presented in Figure 1, the study aims to develop a financing system of nursing care that makes the shift from financing actual nursing interventions and nurse staffing levels to a system that is based on justified nursing interventions and nurse staffing levels.

What nurses do, does not always reflect the needs for nursing care. This is a difficult and long lasting issue. In the pilot study on NMDS [2], nursing activities were measured on three moments in time: (1) 24h in advance (planning), (2) what actually should have been done, (3) what actually was done. There was no difference reported between (2) and (3). Nurses seem to perform the activities that they see as being needed for patients. It doesn't mean however that these activities should have been performed or that other activities shouldn't have been more appropriate. This shows it is of limited avail to ask nurses what interventions would be required if they would have the necessary time and resources.

To investigate which nursing interventions are justified, an evidence-based nursing (EBN) approach will be followed. Therefore the application of evidence based methodology in determining appropriate care is tested for a selection of nursing interventions. A rigorous framework is followed to find, assess and summarize the evidence. The availability of nursing care evidence for a selected number of nursing interventions in the B-NMDS-II is assessed. This implies not only that sufficient research should have been done about a specific nursing intervention, but also that this research should provide evidence of a sufficiently high level.

Nurse staffing is responsible for a considerable portion of the nursing budget and the intensity of nursing activities is an important driver of nurse staffing levels. Therefore, nurse staffing levels are used as a proxy for nurse costing. To meet this objective, quantitative analysis based on empirical data combined with a qualitative approach is pursued. The current nurse financing scheme is mainly based on actual average nurse staffing levels, which means that there is a risk of what is sometimes called as the Matteüs-effect in which the rich gets richer and the poor gets poorer.

For setting justified nurse staffing levels, a Delphi approach with nurse professionals will be followed. The Delphi-approach is agreeing on a standard of what should be a justified nurse staffing level for safe qualitative care. The Delphi method provides an opportunity for experts (panellists) to communicate their opinions and knowledge anonymously about a complex problem, to see how their evaluation of the issue aligns with others, and to change their opinions, if desired, after reconsideration of the findings of the group's work [3].

The final result will be a two-step nurse costing model: (1) a list of nursing interventions and their relative weights that contribute to nursing time needs as a proxy for costs, (2) a model

will be build in linking and explaining nursing costs with Diagnosis Related Groups (DRGs) and other related variables. A major problem in linking both datasets is the different design of data-collection. The B-HDDS is a summary of the hospital stay, collected for all hospitalized patients at time of discharge. The B-NMDS uses a cross-sectional data collection method for a balanced sample of inpatient days. As a consequence, not all DRG's have sufficient nursing data and not all nursing data are representative for the stay of a DRG.

Different scenarios will be explored in order to evaluate the NMDS-II for its use in hospital financing. When a representative sample of Belgian hospital data would be available, these scenarios could be evaluated in comparison with the actual hospital financing rules and data.

The research project, including patient record review, was approved by the Medical Ethics Committee, University Hospital – Catholic University Leuven. All participating hospitals provided a written informed consent. The gathered patient data were treated anonymously and considered as confidential material.

Figure 2 presents the structure of the report as systematically presented in the subsequent chapters. Chapter 1 treats the latest research findings and issues in the construction of financing systems for nursing care. The relationship with DRGs, variability in nursing intensity and the issue of cost compression are discussed. Chapter 2 presents the current financing system for hospital nursing care in Belgium. Strengths and weaknesses of the system are highlighted. Chapter 3 draws on an extensive review of nurse financing systems abroad. The systems of a wide array of countries are presented within a clear taxonomy.

The availability of sufficient evidence based grounding of nursing interventions is the subject of the fourth chapter. It can be considered as a first condition and a starting point to investigate a potential incorporation of justification of activities within nurse financing. Taking this route one step further, chapter 5 expands on the evidence based recommendations as a result in chapter 4, to confront EB knowledge with nationally registered clinical data concerning nursing intervention execution and its indications. This results in a rule set assessing the level of under and over care. In this way the practical implementation of a 'justified care' concept is tested to build on in future financing applications.

The main second aim of the study, the transfer of actual towards nurse staffing needs as a basis for financing, is discussed in chapter 6 and 7. Both a patient case level and a nursing intervention level are tested as alternative approaches. A relative nurse time weighting system of nursing interventions is described. Reliability and validity of the methods used are discussed.

Chapter 8 refocuses on the 'justified care' concept. It assesses the effect in nurse staffing needs between rigorous evidence based care versus actually delivered care. Finally, chapter 9 presents a nurse time needs model as a basis for hospital financing. All results are integrated in chapter 10: General discussion and conclusions.

Figure 2: Structure of the report as an overview of the subsequent chapters

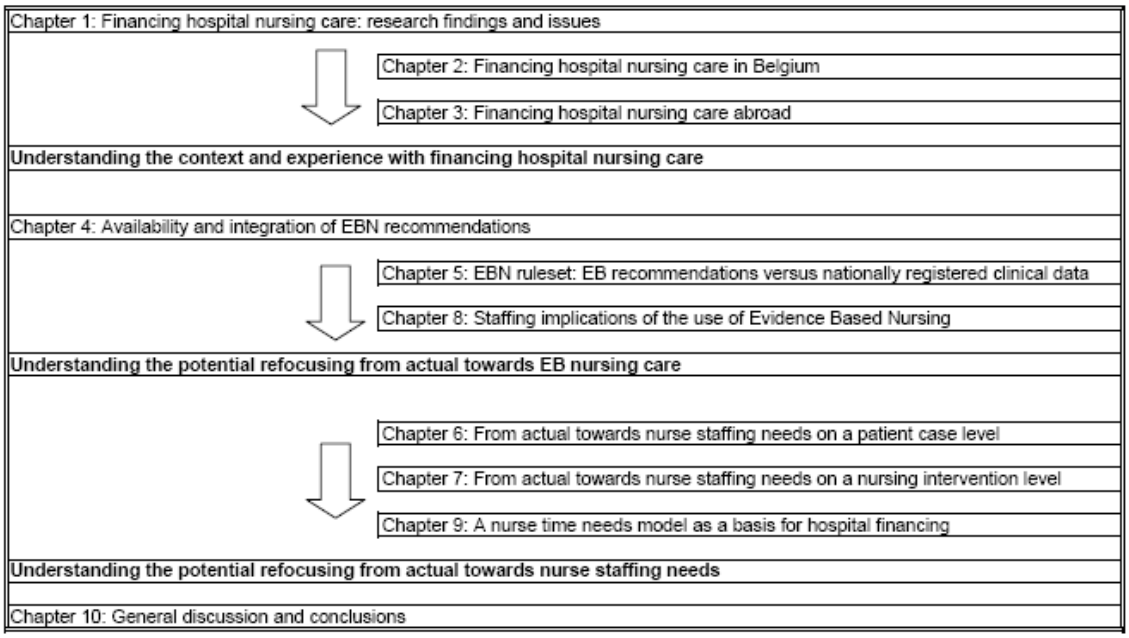


Table I presents a glossary of abbreviations used in the report.

Table 1: Glossary of abbreviations

ADL	Activities of daily living
AEP	Appropriateness Evaluation Protocols
ANA	American Nurses' Association
AP	All Patient
APR	All Patient Refined
AR	Australian Refined
CHF	Chronic Heart Failure
CIHI	Canadian Institute for Health Information
CMG	Case Mix Groups
CMS	Centres for Medicare and Medicaid Services
COPD	Chronic Obstructive Pulmonary Disease
CTI/CIV	Unique Hospital Identification code
DAD	Discharge Abstract Database
DAGS	Danish Ambulatory Grouping System
DBC	Diagnosis Treatment Combinations
DRG	Diagnosis Related Groups
EB	Evidence Based
EBN	Evidence Based Nursing
EBP	Evidence Based Practice
EFN	European Federation of Nurses
EU	European Union
FTE	Full Time Equivalents
GHM	Groupes Homogènes de Malades
GLM	Generalized Linear Modeling

H	Hubert's Mean
HCFA	Health Care Financing Administration
HDDS	Hospital Discharge Dataset
HRG	Healthcare Resource Group
ICD	International Classification of Diseases
ICU	Intensive Care Unit
IGIF	Institute of Financial and Information Management
IMIA - NI	International Medical Informatics Association – Nursing Informatics
IOM	Institute Of Medicine
LEP	Leistungserfassung in der Pflege
LOS	Lenght Of Stay
MKG/RCM	Minimal Clinical Data (see HDDS)
MRSA	Methicillin Resistant Staphylococcus Aureus
MZG	Minimal Hospital Dataset
NHCDC	National Hospital Cost Data Collection
NHS	National Health Service
NIC	Nursing Intervention Classification
NMDS	Nursing Minimum Dataset
NRG	Nursing Related Group
NRS	Numerical Rating Scale
OAG	Oral Assessment Guide
OCDM	Ontario Cost Distribution Methodology
PCSI	Patient Classification Systems International
PFME	Pelvic Floor Muscle Exercise
PMSI	Programme de Médicalisation du Système d'Information
PPS	Prospective Payment System
PRN	Programme Recherche Nursing
RIM	Resource Intensity Measures
RIW	Resource Intensity Weights
ROM	Risk Of Mortality
SD	Standard Deviation
SOI	Severity Of Illness
SPG	Swiss Payment Groups
TIA	Transient Ischemic Attack
UI	Urinary Incontinence
VAS	Visual Analogue Scale
VRS	Verbal Rating Scale
WMS	Workload Measurement Systems
ZIP/ZAP	Zone intensive profile/zone general profile

1 FINANCING FOR NURSING CARE

Costs of nursing staff account for approximately 50% of the total personnel budget and 20 to 30% of the hospital running costs [4]. Although there is high impact on the hospital budget, little is known about the actual relationship between cost of nursing and reimbursement for nursing care.

In most countries, nursing costs are just part of the room and board costs of a hospital. It means that total nursing costs are calculated, divided by the total number of inpatient days, into an average nursing cost per day. As most countries have moved towards a prospective payment mechanism for reimbursing hospital care in which primary tool for reimbursement of hospital care are Diagnosis Related Groups (DRGs), length-of-stay is being used as a proxy for nursing costs. [5]

DRGs is a patient classification system that provides a means of relating the type of patients a hospital treats (i.e., its case mix) to the costs incurred by the hospital [6]. It is a method to group patients in a 'manageable number of groups' on the basis of their economic and clinical homogeneity. Clinical homogeneity is achieved on the basis of agreement in medical diagnosis, co-morbidities, medical procedures, complications. Economic homogeneity is achieved by using first of all the length of stay and later the complete cost of hospitalisation as classification criterion. DRGs have been developed in the United States in the seventies. There is a large family of DRG-systems [7]. It all goes back tot the Yale DRGs in the seventies. Out of these grew the DRGs from the Health Care Financing Administration (HCFA) for hospital payment for Medicare beneficiaries (or more recently called CMS-DRGs according to changing name of HCFA in Centres for Medicare and Medicaid Services (CMS)), the All Patient DRGs (AP-DRGs) which is an expansion of the basic DRGs to be more representative of non-Medicare populations; the All Patient Refined DRGs (APR-DRG) which is incorporating severity of illness and risk of mortality subclasses into the DRGs. Several countries have adapted DRGs into their own DRG-classification, such as Australia (AN-DRGs / AR-DRGs), the Scandinavian countries (NordDRGs, DkDRGs), Germany (G-DRGs), France (GHM), UK (HRG), The Netherlands (DBC), Austria (LDF) etc… In the original version of 1979 the DRG system included 383 groups. In the most recent version the number of groups has been increased to more than one thousand. Belgium is using the APR-DRG version 15.

The original objective of the DRGs was to develop a patient classification system that related the types of patients treated to the resources they consumed. The HCFA DRGs and the AP-DRGs have remained focused on this limited objective. As the health care industry has evolved there has been increased demand for a patient classification system that can be used for applications beyond resource use, cost and payment [5]. Examples of these new objectives are the comparison of hospitals across a wide range of resource and outcome measures, the evaluation of differences in inpatient mortality rates, the implementation and support of clinical pathways, the identification of continuous quality improvement projects, the basis of internal management and planning systems etc.

As soon as DRGs were introduced, most nursing research revealed that DRGs are not very homogeneous to nursing care [8, 9, 10, 11]. This has been measured in different ways. DRGs only explain 20% to 40% in the variability of nursing care. Coefficients of variation for nursing care per DRG are reported varying from 0.22 tot 2.56 [12, 13, 14, 15]. Some DRGs are more nurse intensive as part of nursing in total charges varying from 6% to 25% [14, 16, 17, 18, 19, 20, 21, 22]. The most common critique is that the hospital product is predominantly defined from the medical condition and that nursing is only a cost factor reflected in 'intensity of nursing care' and a measure of how many nurses and minutes that are needed. Welton & Halloran (2005)[23] show that completing DRG-data with of nursing data can improve the prediction for total hospital length-of-stay, total ICU-days and total charges with about 30%.

Outcome Variable	Comparison	$1 - R^2/R^2$	Improved
Hospital length of stay	NDX + DRG vs. DRG	1 − 0.329/0.253	30.0%
	NDX + APR-DRG vs. APR-DRG	1 − 0.335/0.259	29.3%
Intensive care days	NDX + DRG vs. DRG	1 − 0.321/0.186	72.5%
	NDX + APR-DRG vs. APR-DRG	1 − 0.448/0.349	28.3%
Hospital charges	NDX + DRG vs. DRG	1 − 0.372/0.265	40.3%
	NDX + APR-DRG vs. APR-DRG	1 − 0.417/0.327	27.5%
Hospital death	NDX + DRG vs. DRG	1 − 0.637/0.295	115.9%
	NDX + APR-DRG vs. APR-DRG	1 − 0.626/0.254	146.4%
Discharge to nursing home	NDX + DRG vs. DRG	1 − 0.626/0.254	146.4%
	NDX + APR-DRG vs. APR-DRG	1 − 0.406/0.211	92.4%

Figure 3: Improvement in explained variance by adding nursing care information[23]

Although nursing care is not seen very homogeneous within and between DRGs, in most countries nursing care costs are not directly influencing the reimbursement scheme.

John Thompson, a nurse and a member of the Yale University team that devised the DRG, originally proposed to account for hospital nursing care costs and use nursing intensity to adjust the prospective payment to hospitals [24]. Efforts during the 1980's to construct a nursing data model that could be used in the hospital billing and discharge abstract resulted in the proposal of a nursing minimum data set (NMDS) to be included in the hospital discharge and billing abstract [25, 26]. Nursing intensity was one of four nursing indicators along with nursing diagnoses, interventions, and outcomes. Work towards using nursing intensity to account for variability in nursing care essentially ended unfortunately with the death of Thompson in 1992.

Thompson & Diers (1991)[27] report that, already right from the start in 1980, there was a specific requirement from the New Jersey State Nurses' Association (NJSNA) to allocate nursing intensity to DRGs. The provision was in part responsible for the NJSNA's support for the DRG implementation. DRGs would be "weighted" by a nursing intensity factor that would be reflected in reimbursement rates. The study produced an instrument to measure nursing resource use: Resource Intensity Measures (RIMs) [28]. Because of a flawed methodology, RIMs were never implemented as part of the New Jersey scheme.

In 1985, the Health Care Financing Administration (HCFA) made available contract funds through the American Nurses' Association (ANA) to begin an investigation of nursing intensity within DRGs. A number of other small studies using DRGs were conducted independently. In general, all of the studies focused upon the extent to which DRG assignment did or did not predict nursing requirements within DRGs. There were many methodological limitations involved in these studies. Only a selection of DRGs was investigated. But most of the studies showed that DRGs were not very homogeneous to nursing care. In 1987, the Yale Health Systems Management group, lead by B. Fetter and J. Thompson started a study to develop and test models of accounting for nursing resources within DRGs. The study took place in 5 hospitals, including 139498 patient records. Based on the analysis, a relationship between total nursing time and length-of-stay is consistent and regular (Figure 4).

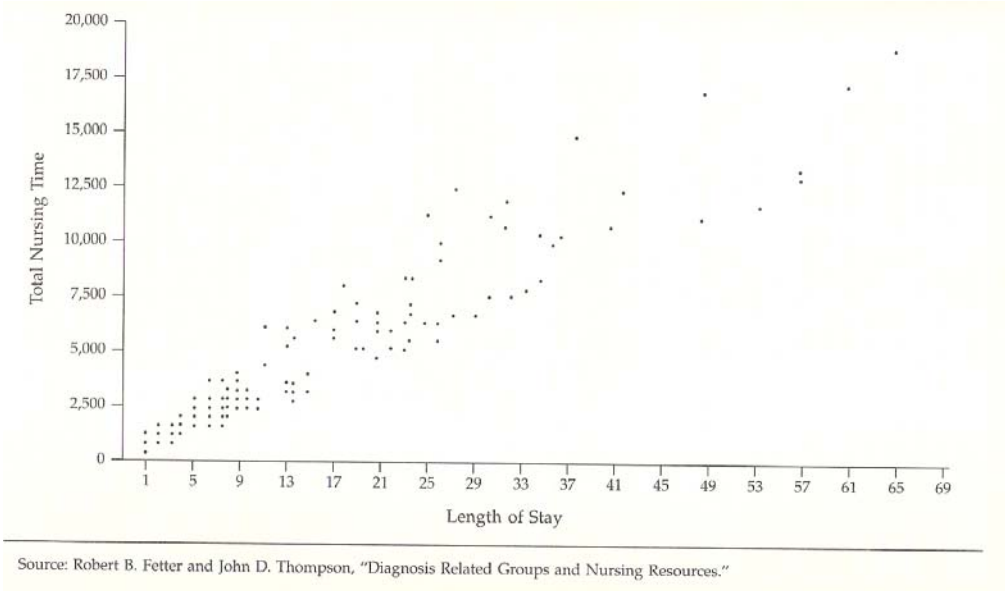


Figure 4: Total nursing time by length-of-stay, non-icu, DRG014, specific vascular disorders except TIA, hospital A [27]

Based on this relationship, a model of relative nursing intensity could be constructed by regressing nursing time on length-of-stay to produce the beta weight or slope of the line (Figure 5). Line A would represent a DRG with relatively low requirements for nursing over rather long lengths of stay. Line D would represent a DRG with relatively high requirements for nursing care over shorter stays. The beta weights can be interpreted as the increment in nursing time per day of stay. "Influential observations" identified by Cook's D statistic were eliminated from further analysis. The correlations of these beta weights among the 5 hospitals varied from 0.34 to 0.62.

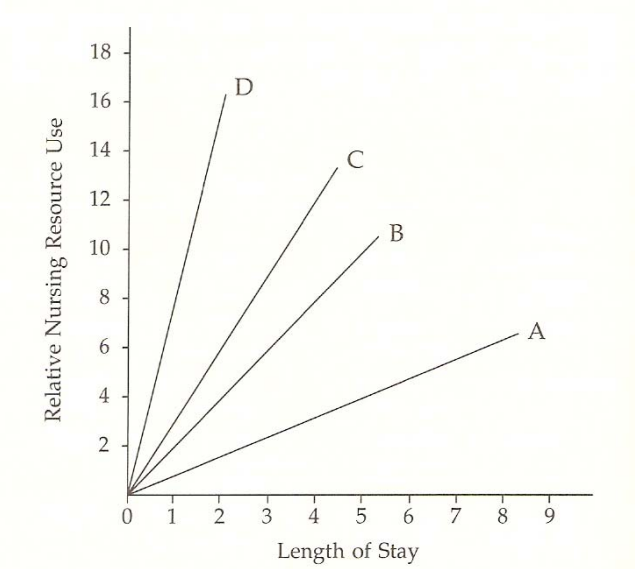


Figure 5: Graphic representation of nursing allocation statistic [27]

In the Yale study five patterns of nursing intensity across days of stay were identified:

- Elective surgical patients show a peak of nursing intensity on the second day of stay, the first postoperative day
- The pattern of trauma patients begins very high and then slopes down toward the end of stay, but never reaches the lowest levels since that patients often retain disabilities requiring nursing attention
- Patients in a terminal phase of illness often show a pattern that begins with low level of nursing intensity increasing as the illness progress
- Some patients show essentially a flat pattern
- Other patients show no pattern at all (e.g. chronic elderly patients)

In an effort to simplify the model, a panel of nurse clinicians was asked to group DRGs that which would require similar amounts of nursing care on routine floors. Six clusters were created along the two dimensions of risk and dependency. It was assumed that patients were admitted to hospitals for nursing care either because they cannot care for themselves or because they need inpatient monitoring and treatment. The analyses suggest that there is relative consistency across hospitals in routine care nursing resources consumed by patients in particular DRGs, even if the actual minutes or beta weights differ. The assigned nursing time per cluster varied from 210 minutes in cluster 1 to 450 minutes in cluster 6. Further analysis on the level of DRGs between he different hospitals revealed that there is a high variability in nursing time from DRG to DRG and from hospital to hospital. The six groups were further refined and tested by Diers & Bozzo (1997)[29].

"The comparisons among hospitals suggested that there is so much yet to be understood about the difference in nursing intensity that any change in the reimbursement policy, such as weighting DRGs by nursing intensity, would be premature" [27].

Out of the work of John Thompson, two separate tracks were developed. The first track is a more professional track leading to the development of the concept of the minimum nursing datasets. The second track is on workload measurement.

The first track on nursing minimum datasets (NMDS) was seen as a continuation of the work was initiated by Florence Nightingale in her "Notes on Hospitals" (1863)[30]. Following initiatives on hospital discharge datasets in the seventies, from 1977 a 'Nursing Minimum Data Set' was prepared in the U.S.A. [25]. A proposal for a NMDS was finalized in 1985. A NMDS was defined as "a minimum data set of items of information with uniform definitions and categories concerning the specific dimension of nursing that meets the information needs of multiple data users in the health care system" [25]. The dataset was mainly structured according to nursing diagnosis, nursing intervention and nursing results of care. A fourth element was included (as suggested by John Thompson): intensity of nursing care. It was defined as the total number of hours of nursing care per individual patient. Except for some limited use in collecting some data for research purposes, the dataset has never been implemented in a systematic way [31].

At the same time a NMDS was developed in Belgium especially focusing on a limited number of nursing interventions (23), mandated by the Belgian Ministry of Public Health. The development resulted in a systematic collection of nursing data in Belgian hospitals since 1988 [3].

Both models (USA, Belgium) inspired different countries to develop their own nursing minimum dataset such as Canada [32], The Netherlands [33], Finland [34], Switzerland [35], Portugal [36], Sweden [37] and Ireland [38]. In most countries the development was tested on a small scale for research purposes. No systematic data collection has started yet.

An initiative has started in 1997 within the International Medical Informatics Association – Nursing Informatics (IMIA-NI) to develop an International Nursing Minimum Dataset (I-NMDS) [31]. The work has been piloted in 2004 in comparing US, Belgian and Swiss data. Although the need for comparable data has been high according to the ICN agenda and strategies [39], no final I-NMDS structure has been defined yet.

The second track on workload measurement was focused in using different (mainly existing) patient classification systems for measuring nursing workload. The common denominator here is nursing time. Many examples can be given such as France, Australia, and Switzerland.

In France, the framework of the 'Programme de Médicalisation du Système d'Information (P.M.S.I.)' was launched in 1987 carrying out out experiments with the collection of data on the intensity of nursing care. Proposals have been made to complement this with data concerning the nature of nursing care [40].

In Australia, the Australian Nursing Association, lead by Picone a research project to develop cost weight per DRG [41]. The PAIS patient classification system developed by Hovenga was used to sort DRGs in similar groups.

Similar work was done in Switzerland in developing the LEP patient classification system and linking it to DRGs.

Both tracks were confronted with serious problems in linking nursing data with DRGs. A first issue in the NMDS-track is that definitions were set and unequivocal, but no data were available. In the workload-track, many hospitals had data available from patient classification systems that are used internally for workload measurement. But often they were not representative or comparable.

Both approaches also lead to different ways in linking nursing care and DRGs (Figure 6), as has been suggested by Sermeus et al (2006)[5] and Fischer (2002)[7]. The first approach leads to redefining the "patient product" by linking DRG and nursing care data. The second approach leads to describing the "nurse costing variable".

For the first approach in describing the patient product, three methods can be followed.

- The first method is that the DRG-classification is kept intact and is solely built upon medical data such as medical diagnoses and interventions and some patient data such as age, etc. Nursing is not involved in describing the patient product.
- The second method is that the basic structure of DRGs is kept intact, but that complications and co-morbidities are not only refined on medical data only but supplementary on nursing diagnoses and interventions.
- The third model is that independent from the DRG-classification a nursing classification, so called Nursing Related Group (NRG) could be developed.

The second approach is describing the nurse costing variable.

- The most frequently used method is to fix nursing costs as the total sum of nursing hours divided by the number of inpatient days, giving a measure of "nursing hours per patient day". Variability in nursing care is only measured by the variability in length of stay. For this method no nursing data are required.
- A second method recognizes that nursing care differ between DRGs. This leads to fixed (relative) nursing cost weights per DRG, which are calculated by linking

nursing workload systems to DRGs. For this method, DRGs need to be calibrated for nursing care. Sample data can be sufficient.

- The third method is describing the variability of nursing care within DRGs. For this method, continuous nursing data collection is required.

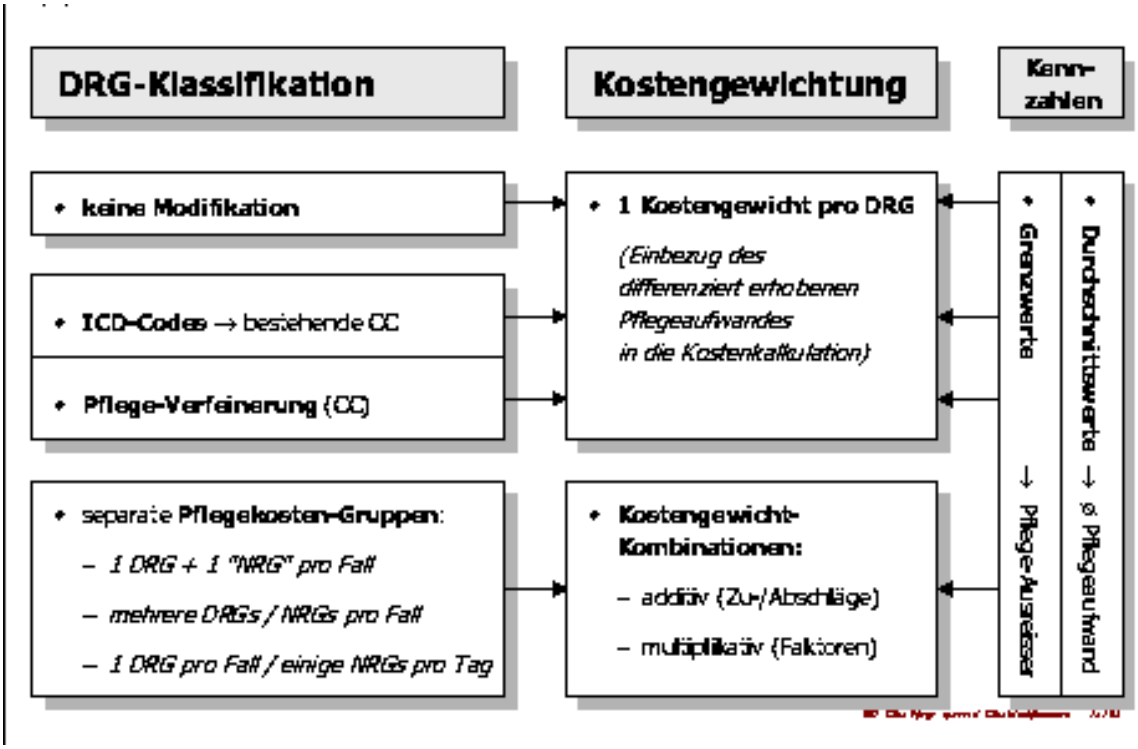


Figure 6: Linking DRGs and nursing care (http://www.fischer-zim.ch/)[7]

The NMDS-track is more focusing on the first approach. The workload-track is more focusing on the second approach. Because of the fact that the workload-track put nursing more as a cost factor instead of a production factor, this approach has been often disapproved by the nursing profession [39].

Characteristics	*Professional track by NMDS*	*Managerial track by workload measurement*
Focus	*Products*	*Costs*
Common denominator	*Added value*	*Nursing time*
Approach	*Uniform definitions and data collection*	*Making use of existing data*
Link to hospital financing	*Adjusting DRGs*	*Determining cost-weights*

The challenge is in bringing these two tracks together. There is no value in having a well-defined nursing minimum data sets and having no data to link with. There is no value in having a lot of intra-hospital workload without overall comparability. Workload and nursing data need to go hand in hand. There is no use in discussing a workload figures without information on the patient problem, the nursing intervention or patient outcomes. There is no need to develop nursing classifications aside from DRGs. There is a high need for integration. The aim of this study is to develop a sound model to link both tracks in providing uniform definitions through the B-NMDS-II, collecting the data, deriving nursing workload data and linking them with DRGs.

2 FINANCING HOSPITAL NURSING CARE IN BELGIUM

The financing of hospital care in Belgium is a federal matter. The federal ministry of health is actually using a mixed system for the financing of its 116 acute care (non psychiatric) hospitals (42 public, 74 private); of which 7 are university hospitals. The global financing of hospital care can roughly be divided in [42]:

- one part (40%) is based on a prospective based system;
- a second part (40%) is based on a payment per medical activity (nomenclature);
- a third part (10-15%) consists of the budget for drugs;
- a fourth part (5%) is based on individual agreements between hospital and the federal institute for sickness and invalidity insurance;
- a fifth small part (2%) is paid directly by the patient.

In the prospective payment system the APR-DRG system (grouper ICD-9-CM version 15 from 3M) is used to define the hospital's case mix product based on ICD-9-CM coding system.

Per APR-DRG, severity of illness and age category the federal government defines a federal Expected Length of Stay (ELoS) based on the hospital discharge dataset called "Minimale klinische Gegevens / Résumé Clinique Minimal MKG/RCM" which captures patient characteristics, diagnoses, interventions, hospital length of stay. In order to have a stabile ELoS a reference period of three years is used. Based on several algorithms inliers and outliers are defined. For inliers the national average length-of-stay, called justified length-of-stay is used financing hospital care. For outliers, different financing rule are defined.

Multiplying the ELoS by the number of admissions per APR-DRG, severity of illness and age category and per hospital gives the number of expected patient days per hospital. These patient days are translated into a number of "justified beds" by dividing by 365 and a normative occupation rate which is 70% for paediatrics and maternal wards; 90% for geriatric wards and 80% for surgical, medicine wards. This measure defines the volume component of the hospital's budget.

The second part is the cost part of the hospital budget. The budget part is divided in three parts: A (capital costs), B (operational costs), C (corrective measures). Part B is divided in different cost centres: B1 (general operational costs (administration, maintenance, laundry, …), B2 (clinical costs), B3 (medical-technical departments), B4 (fixed clinical costs), B5 (pharmacy costs), B6 (extra-legal financial benefits), B7 (added costs for teaching hospitals), B8 (social costs), B9 (extra-legal financial benefits).

Part B is 88% of the hospital budget of which B2 is 48% and B1 is 27%. B2 is holding the budget for all clinical cost centres such as nursing wards, operating theatre, and emergency care. The budget is for nursing staff and medical supplies.

The budget has two parts: a fixed budget and a variable budget. The fixed budget is calculated directly on the number of justified beds. The point of departure is minimal nurse staffing ratios that have been set in the past for various types of nursing wards (table 2).

To calculate the budget, a point system is used. Every year, the total prospective budget for hospitals is approved by the Council of Ministers. This budget is divided by the number of financial B2-points that are earned by the hospitals. The result is a financial value of a point, what allows to calculate the final budget for each hospital and to stay within the provisional budget limits. In 2006, the financial value of a B2-point was 20205 Euro.

Table 2: FTE formation by index

Nursing Ward	Nurse staffing / justified beds	B2- points per justified bed
Internal Medicine (CD)	12/30	1
Surgery	12/30	1
Paediatrics (E)	13/30	1
General hospitalization (H)	9/30	0,88
Maternity (M)	14/24	1,46
Maternal Intensive Care (MIC)	1,50/1	3,75
Neonatal Intensive Care (NIC)	2,50/1	6,25
Geriatrics (G)	13,33/24	1,36
Intensive Care (I)	2,00/1	5
Psychiatry Acute Care (A)	16/30	1,33
Child Psychiatry (K)	16/20	2

For some nursing wards, there is also a variable part in the budget. This is calculated on several criteria.

For Surgery and Internal Medicine, two types of criteria are taking into account:

1. an average cost-weight based on medical and surgical interventions per bed (according to a federal nomenclature of interventions) (20% weight)
2. an average cost-weight based on B-NMDS-I per patient per day (80% weight)

Hospitals are ranked according to these weights in deciles (groups of 10% hospitals). The hospitals that are in the lowest ranked group (decile 1) don't get any variable budget. The hospitals ranked in the highest ranked group (decile 10) get most additional budget. Table 3 gives an overview of these upper and lower limits per decile and additional B2-points for surgery and internal medicine wards. The use of the deciles system implies that each hospital is reimbursed as a function of its relative position compared to other hospitals and not based on its actual or aspired performance. The use of deciles in this ranking is arbitrary. As a side effect of this type of ranking, small differences between hospitals can give lead to big differences in reimbursement [42]. And big existing differences between hospitals within the same decile aren't taken into account.

An example can help to explain how the calculation goes. Suppose that in a given hospital A, the nursing wards on surgery and internal medicine have a medical nomenclature cost-weight of 75 and a nursing NMDS-cost-weight of 0, 53. The medical cost-weight of 75 is ranked in decile 7 (0, 14 points). The nursing cost-weight is ranked in decile 5 (0, 06 points). The final result is that the nursing wards will get 0,076 additional B2-points per justified bed (0, 14*20% + 0, 06*80%). Given the 1 point per justified bed in the fixed part of the budget, the budget for surgery and internal medicine equals 1,076 B2-points per justified bed. Suppose that there are 100 beds justified. 100 beds equals 107, 6 B2-points, multiplied with 20205 Euro give a budget of 2,174,058 Euro.

Table 3: Hospital supplementary financing parameters for surgery and internal medicine

Decile	Medical nomenclature cost	Nursing B-NMDS-I cost	Additional B2-

	weight (20%)		weight (80%)		point per justified bed
	lower limit	upper limit	lower limit	upper limit	
1	0	59.21	0	0.48354	0
2	59.52	63.39	0.48419	0.50135	0
3	63.51	65.23	0.50162	0.51339	0
4	65.52	67.45	0.51673	0.52335	0.02
5	67.65	69.98	0.52407	0.5402	0.06
6	70.2	73.31	0.54215	0.54891	0.09
7	73.83	77.39	0.55092	0.56729	0.14
8	77.68	89.28	0.56789	0.5823	0.18
9	89.63	94.4	0.58255	0.60627	0.27
10	95.16	111.15	0.60628	0.68516	0.34

A similar calculation is made for the paediatrics wards. The main difference is that the medical nomenclature data weight for 70% and the nursing NMDS data weight for 30% (table 4)

Table 4: Hospital supplementary financing parameters for paediatrics

Decile	Medical nomenclature cost weight (20%)		Nursing B-NMDS-I cost weight (80%)		Additional B2-point per justified bed
	lower limit	upper limit	lower limit	upper limit	
1	0	55.84	0	0.39856	0
2	56.52	59.91	0.40052	0.41472	0.01
3	60.29	63.49	0.41528	0.42786	0.05
4	64.67	67.47	0.42903	0.44215	0.10
5	67.85	70.71	0.44217	0.45645	0.13
6	70.78	73.20	0.45666	0.47153	0.15
7	73.94	76.73	0.47357	0.49345	0.18
8	77.06	83.95	0.49922	0.52983	0.20
9	84.63	90.05	0.53210	0.57186	0.25
10	91.12	135.85	0.59297	0.70440	0.38

For intensive care units, the calculation is more complex. The first difference is that deciles from 7 to 10 are divided into smaller groups, what make the additional point curve even steeper. Intensive care beds are calculated on top of surgical, internal medicine and paediatrics wards as percentage of justified beds for intensive care or as supplementary budgets for the routine care wards. Three criteria are used for the calculation of additional B2-points.

- A selection of medical interventions from the nomenclature which are characteristic for intensive care such as resuscitation, artificial ventilation, invasive monitoring etc. Ten medical interventions are in this list. A medical nomenclature cost weight is calculated based on average charges for these interventions on surgical, internal medicine, geriatrics and paediatrics wards) (weight is 20%)
- The number of inpatients days with an intensive care profile measured by the NMDS. Within the B-NMDS, 5 intensive care (ZIP) and 23 non-intensive care (ZAP) profiles have been identified in which all inpatient days on surgery and internal medicine are classified. The ratio of ZIP-days to ZAP-days is taken as a measure of intensity (weight is 40%)
- An ICU-case mix measure (NPercIZ) based on the expected number of days on ICU per APR-DRG and severity of illness (weight is 40%)

Table 5: Hospital supplementary financing parameters for intensive care

Decile	Medical nomenclature cost weight (20%)		ZIP/ZAP-NMDS-cost weight (40%)		ICU casemix measure 40%)		Additional B2-point per justified bed	% beds intensive care
	lower limit	upper limit	lower limit	upper limit	lower limit	upper limit		
1	0	2.46	0	0.03561	3.5190	3.9520	0.08	2%
2	2.49	2.79	0.03753	0.04608	4.1100	4.7010	0.08	2%
3	2.81	3.40	0.04665	0.05459	4.7040	5.0430	0.08	2%
4	3.41	3.72	0.05530	0.06185	5.0480	5.4180	0.08	2%
5	3.76	4.15	0.06218	0.06641	5.6190	5.9400	0.10	2.5%
6	4.16	4.74	0.06669	0.07573	5.9470	6.2460	0.13	3.25%
7 lower	4.79	4.94	0.07594	0.07855	6.2630	6.7340	0.15	3.75%
7 upper	4.98	5.09	0.07939	0.08295	6.7540	6.9110	0.17	4.25%
8 lower	5.10	5.54	0.08337	0.08824	7.0270	7.2560	0.20	5%
8 upper	5.57	6.00	0.09069	0.09458	7.5080	7.9030	0.21	5.25%
9 lower	6.01	6.32	0.09617	0.09819	7.9120	8.3480	0.24	6%
9 upper	6.37	6.76	0.09898	0.10502	8.5510	8.8280	0.28	7%
10 lower	6.94	7.24	0.11419	0.12484	9.0410	11.7240	0.33	8.25%
10 upper	7.93	9.69					0.41	10.25%

2.1.1 How is "nursing care" measured to be included in the Belgian hospital financing scheme?

Belgium is one of the few countries that complement this HDDS with a nationwide uniform Nursing Minimum Dataset (NMDS) for a balanced sample of yearly 20 inpatient days since 1988. The mandatory registration resulted in an extensive dataset of more than 15 million selected in-patient days for some 6 million selected patients in all 2.500 nursing units in all Belgian hospitals. Since 1994, the B-NMDS is used in the hospital financing system. At the

start, the impact of the hospital budget was limited. In 2006, about 5, 5% of the hospital budget is determined by the B-NMDS.

The B-NMDS-I consists of 23 nursing interventions (see table 6)

Table 6: B-NMDS-I nursing interventions

care relating to hygiene (no assistance, supportive assistance, partial assistance, complete assistance)	*care relating to mobility (no assistance, supportive assistance, partial assistance, complete assistance)*
care relating to elimination (no assistance, supportive assistance, partial assistance, complete assistance)	*care relating to feeding (no assistance, supportive assistance, partial assistance, complete assistance)*
tube feeding (yes, no)	*special care on mouth (frequency/24h)*
decubitus preventive care (frequency/24h)	*assistance in getting dressed (yes, no)*
care of patient with tracheotomy or endotracheal tube (artificial ventilation or not)	*nursing anamnesis report (yes, no)*
discharge teaching to patients (occasionally, according to programme)	*emotional support (yes, no)*
supervision to mentally disturbed patient (passively - reality orientation training)	*isolation for preventing contamination (yes, no)*
monitoring of vital signs (highest frequency/24h)	*monitoring of clinical signs (highest frequency/24h)*
attending on traction, cast, external fixator (yes, no)	*drawing of blood specimen (frequency/24h)*
administration of medication (intramuscular, subcutaneous, intradermal) (number of doses/24h)	*administration of medication (intravenous) (number of doses/24h)*
attending on continuous infusion (number of lines)	*surgical wound care (number of interventions/24h)*
traumatic wound care (surface + number of interventions/24h)	

Based on the B-NMDS, all inpatient days are classified according to 28 different zones, called the NMDS-map. The "NMDS-map" characterises every nursing unit with respect to the other nursing units. In this two-dimensional space the horizontal axis refers to continuum of intensity of care, while the vertical axis refers to the balance of technical nursing care versus basic nursing care (Figure 7). Some zones (19, 20, 24, 25, and 28) are labelled as intensive care zones (ZIP). All other zones are labelled as non-intensive care (ZAP). Every zone is attributed a weight according to the nurse staffing and qualification mix. This weight is based on actual nurse staffing and qualification data from all Belgian hospitals (table 7).

The weighting of these zones is highly questionable [42].

First of all, the definition of these 28 zones hasn't changed since 1992. The cost-weighting per zone is recalculated every time a new reference year was used. It is remarkable that the overall nursing cost-weight has only increased with 2%. It is even more remarkable when we see that there is a shift in inpatient days to more intensive zones (ZIP-zones). The explanation is that some zones are weighted lower than before. In 2000, 18 of the 28 zones weighted lower than in 1998. Ten of these 28 zones weighted even lower than in 1992.

Secondly, these zones are weighted according to the actual nurse staffing data from all nursing wards. Nursing wards are used for the calculation of the weight of a zone if their average NMDS-profile is in this zone. An example can be helpful. Zone 6 (see figure 7) has a profile of patients who are independent for their basic care needs. From a NMDS-perspective, no nursing interventions are performed except some routine monitoring of vital signs. The number of patients in this zone is limited to 1, 4% in 2000. Because of reduction in length-of-stay, this type of care becomes very rare indeed. Even more rare are

nursing ward with this profile as average profile. In 1995 there were 35 nursing wards with this type of profile which could be used for determining the cost-weight of zone 6. Only 3 of these wards were surgical, internal medicine or paediatric wards to which the NMDS is used in determining additional budgets. The majority of wards were acute psychiatric wards. It means that the weight is mainly base on nurse staffing data from psychiatric wards and applied to inpatient days from surgical and internal medicine wards (54, 7%). In 2000, only one nursing ward with an average profile of zone 6 was found. It is a day clinic. The actual nurse staffing of this nursing ward has been used for determining the cost weight of this zone. The result is that for 1, 4% of all inpatient days, the cost weight has been increased with 52%.

Figure 7: NMDS Map. Positioning of nursing wards in 28 care zones, based on NMDS dimensions

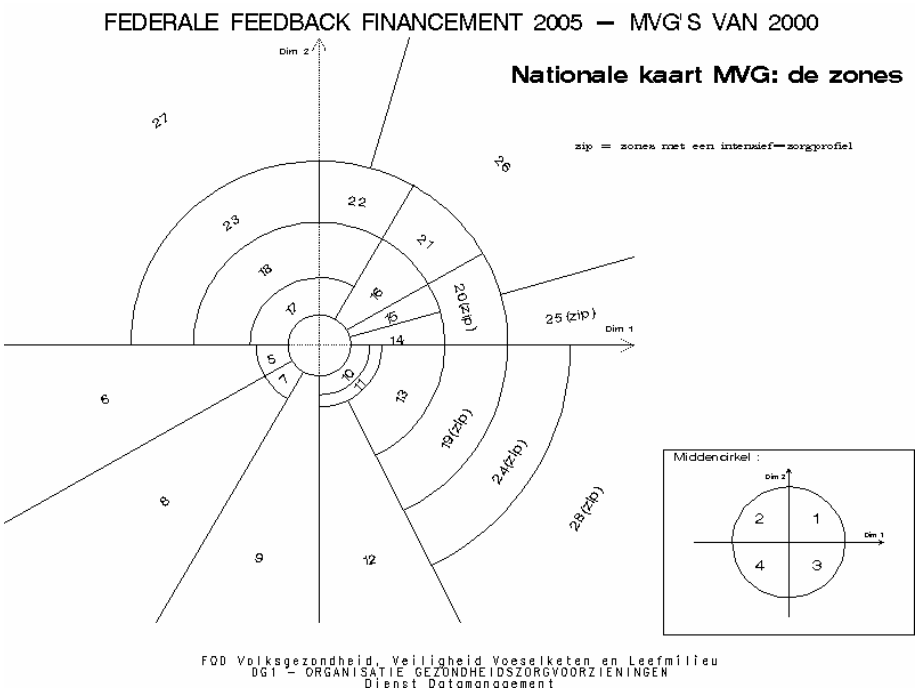


Table 7: Weighting of 28 care zones, based on actual nurse staffing data

Zone	Qualification	FTE / patient day	Weight
1	0.4579	0.3872	0.3794
2	0.4577	0.3328	0.3260
3	0.5664	0.4579	0.4725
4	0.4993	0.3529	0.3528
5	0.4846	0.3337	0.3312
6	0.4166	0.2866	0.2751
7	0.5363	0.3760	0.3825
8	0.5735	0.3208	0.3321
9	0.5545	0.3555	0.3648
10	0.6028	0.6564	0.6888
11	0.6889	0.8652	0.9437
12	0.6571	0.9151	0.9842
13	0.6509	1.1438	1.2268
14	0.6693	0.9132	0.9875
15	0.6877	0.9375	1.0220
16	0.4539	0.4274	0.4179
17	0.4267	0.3471	0.3348
18	0.3742	0.3556	0.3341
19 zip	0.6474	1.3383	1.4331
20 zip	0.6483	1.3353	1.4305
21	0.5952	1.2025	1.2575
22	0.3789	0.3706	0.3490
23	0.3745	0.3554	0.3340
24 zip	0.6641	1.4995	1.6177
25 zip	0.6642	1.4988	1.6171
26	0.5910	1.2010	1.2536
27	0.3842	0.3694	0.3488
28 zip	0.6674	1.5508	1.6755

Discussion of the Belgian nursing care financing system

A major advantage of the Belgian hospital financing system for nursing is that nursing care data are taken into account. It seems obvious, but is not done in many financing systems. There is however room for improvement:

- The system is complex and not transparent, because of the different criteria, cost-weights, deciles according to the type of beds.
- Not all the nursing wards are considered in the system: e.g. geriatric nursing wards are not taken into account when considering the variable part of the budget.
- The NMDS-zones are calibrated with actual staffing data. This means that the historical way of allocating resources is included in the system. The way how these weights are calculated are highly discussable.
- The use of the deciles system means that a global increase in nursing intensity in the Belgian hospitals is not taken into account. The system always relatively redistributes the resources over all the hospitals.
- APR-DRG is only used in the volume component of the budget. It is assumed that "the supplementary part" of the budget is equally distributed over all APR-DRG within a hospital. It means that the impact of changes in case mix is not controlled.
- The major disadvantage is that the impact reduction in length-of-stay by the hospital on nurse staffing is not known, redistributed through the "deciles" system and mainly carried by increased nursing workload.

3 FINANCING HOSPITAL NURSING CARE: LITERATURE REVIEW

3.1 INTRODUCTION

In all countries hospital nursing care is part of the operating costs of a hospital. Some countries adjust the financing system for nursing care. Most of the countries don't and treat nursing costs as part of room and board costs. In some countries, Diagnosis Related Groups (DRGs) are used for reimbursement. Nursing costs are integrated in the cost weight per DRG. In some other countries, nursing costs are treated separately.

The goal of this international review is to understand how the acute care hospitals were financed and more particularly those based on a DRG system.

3.2 RESEARCH QUESTION

The focus of this study part is aimed at the following research questions: Which tools and methods are used abroad in financing hospital nursing care? Which are the characteristics of the methods used?

3.3 METHODOLOGY

Literature on hospital financing systems is not easily found in the scientific literature. The main literature is found in governmental and institutional reports. The search strategy has been adapted to that. Four main sources have been identified:

- Health basket (http://www.ehma.org/projects): The project was led by the European Health Management Association and funded by the 6th EU-Framework Programme [43]. The project was launched in April 2004 and will be complete in March 2007. The main objective of the project is to describe how different countries define the services provided within the system by analysing both the structure and contents of benefit 'catalogues' (or 'baskets'). The participating countries are Denmark, France, Germany, Hungary, Italy, Poland, Spain, The Netherlands, and the United Kingdom. The detailed description and comparison of hospital financing systems in the nine EU-countries has been published in a special issue of the Journal of Health care Financing Science, 2006 [44, 45].
- HOPE [46]: the European Hospital and Healthcare Federation (www.hope.be) did in 2006 a survey among member states about DRGs and hospital financing. The report isn't finalised yet. There was permission to use the preliminary data. Nineteen EU-countries were answering the questionnaire.
- To have more precise information on the financing of nursing in the various hospital financing systems, we performed a survey (APPENDIX 1) among key-representatives of patient classification systems international (PCSI), the international Medical Informatics Association, workgroup on Nursing Informatics (IMIA-NI) and the national representatives of the European Federation of Nurses (EFN). Out of 17 contacts, six replied (Switzerland, USA, Germany, UK, Ireland and the Netherlands).

- A snowball search technique has been used to finding for additional reports and articles, using Medline, websites from governmental bodies, financial and advisory institutes, and professional organisations. Following key-words have been used: DRG, Diagnoses Related Groups, Case-Mix, Hospital Financing, Hospital costs, Nursing Financing, hospital reimbursement system nursing, financing (system) hospital care, prospective payment system.

Sixteen countries will be discussed in more detail: Australia, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Luxembourg, New Zealand, Portugal, Spain, Switzerland, The Netherlands, United Kingdom and the United States.

3.4 RESULTS

The hospital financing systems can be divided in two main types: using DRGs (or a comparable patient grouper system) and not using DRGs. In Europe, almost all countries are using DRGs for hospital financing [46]. The exceptions are Luxemburg, Cyprus, Green and Czech Republic. The Belgian system is somewhat special because it is using DRGs to regulate the number of "justified" inpatients days. The allocation of the final hospital budget is done by different cost centres. The Belgian can be called a mixed system.

The way nursing costs are linked with DRGs can be organised in three different ways [7].

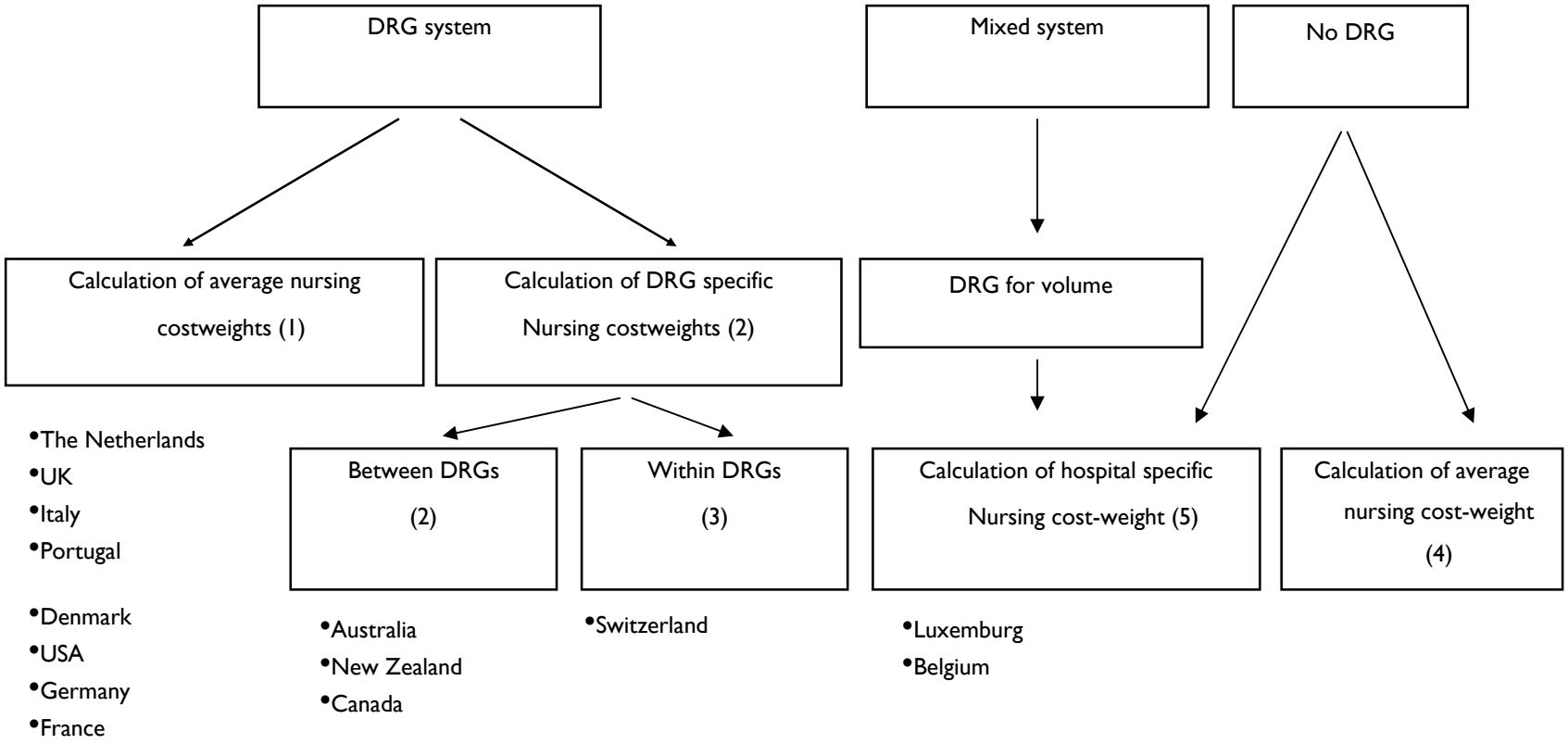
- Calculation of an average nursing cost per patient day. The cost of nursing care is directly related to the number of inpatient days. Nursing costweights per DRG are directly related to length-of-stay.
- Using a DRG specific nursing cost weight. It is done by using nursing workload systems and calculating average nursing time per DRG. Some DRGs are more nursing intensive than other DRGs. Nursing cost weigths differ between DRGs
- Using not a fixed, but a variable nursing cost weight per DRG. It is done by linking DRG and nursing data on a patient level. It leads to a hospital specific nursing cost weight per DRG.

If no DRGs are used, nursing costs can be calculated in two ways:

- Calculation of a average nursing cost per patient day. The cost of nursing care is directly related to the number of inpatient days.
- Calculation of a nursing cost that is related to nursing workload. This results in an hospital specific nursing cost weight.

Some countries have some of these five types of nursing costs calculations in use, have had some experience with these systems but had left them or have some plans to introduce them later. Type 4 is similar to type 1 without DRGs. Type 5 is similar to type 3 without DRGs.

Figure 8 describes the five types on how nursing costs can be calculated within the hospital financing system. The literature review is reported according this structure.


DRG system
Mixed system
No DRG
Calculation of average nursing costweights (1)
Calculation of DRG specific Nursing costweights (2)
DRG for volume
•The Netherlands
•UK
•Italy
•Portugal
Between DRGs (2)
Within DRGs (3)
Calculation of hospital specific Nursing cost-weight (5)
Calculation of average nursing cost-weight (4)
•Denmark
•USA
•Germany
•France
•Australia
•New Zealand
•Canada
•Switzerland
•Luxemburg
•Belgium

3.4.1 Countries using DRGs without adjusting for nursing care

The majority of countries are using DRGs without adjusting for nursing care. A few examples from The Netherlands, UK, Italy and Portugal are discussed. The approach used in most countries is very comparable. DRGs are used to group patients. A uniform costing methodology is used to calculate standard cost per DRG. For allocation this costs, various cost drivers are used. The main cost driver to allocate nursing costs is the number of inpatient days.

In **the Netherlands** a case-mix system based on "diagnosis treatment combinations" (DBCs) for the registration and reimbursement of hospital and medical specialist care was introduced in February 2005 [47]. Currently, there are approximately 29,000 DBCs, as opposed to DRG based systems that mostly consist of approximately 600 to 900 DRGs [48]. To describe an episode of care in the DBC case-mix system, at least 3 dimensions have to be specified: the type of care (regular care, emergency care, and chronic periodical check up), the diagnosis (ICD-10 coding) and the treatment axis (expresses the setting which is either "outpatient" "in day-care" or "with clinical episode" and the nature of the treatment which specifies whether treatment is conservative, surgical, or whether it includes a major non-surgical intervention). The DBC covers the entire treatment episode related to the same diagnosis, including the hospital admission, medical interventions and preceding and subsequent outpatient visits. DBC relies on an episode-based registration within hospitals. One patient can have multiple DBCs at the same time. A distinction is made between DBCs with fixed prices (list A - 95% of DBCs and 92% of budget) and with negotiable prices (list B of elective DBCs - 5% of DBCs and 8% of budget).

During the introductory period of the DBC system, unit costs of DBCs were calculated using information of the front-runner hospitals about the years 2003 and 2004. During this period, all front-runner hospitals adopted a uniform product-costing model to calculate the unit costs of hospital services and DBCs. In the product costing model, a distinction is made into intermediate and final products. The final products in the model are DBCs. Intermediate products are the health care services such as inpatient hospital days, outpatient visits, surgical interventions, etc. The model consists of two parts. The first part involves the calculation of unit costs of intermediate products. In the second part, data about resource use profiles and unit costs of intermediate products are used to calculate the unit costs of DBCs. Hospital departments producing intermediate products are final cost centres. These include inpatient and outpatient clinics, laboratories, operating rooms, radiology departments, etc. Departments not providing patient care are called support cost centres. In the unit cost model, costs of all support cost centres are assigned to final cost centres using direct allocation. For some cost centres such as operating room or radiology, more sophisticated cost allocation drivers have been used. The DBC cost allocation driver for nursing wards is however kept very simple as the number of patient days. Nursing intensity is assumed equal for all days and all patient groups.

The National Health Service (NHS) in **England** is since April 2004 introducing a national cost-per-case tariff system for the reimbursement of hospital services [49]. The new reimbursement system for hospital care, known as Payment by Results (PbR), is being phased in over a four-year period. Inpatient stays (or "spells") and day case activity are priced according to national tariffs for each Healthcare Resource Group (HRG). An HRG is similar to a DRG. The tariff for each spell is set by the Department of Health at the start of each financial year, and is based on the average costs of all NHS hospitals in England. HRGs are seen as "units of currency" within the health service, allowing for costing across services. From 2005/2006 the number of HRGs increased to about 1,000. To prevent proliferation of new HRGs, a new HRG must encompass at least 600 cases nationally and incur over £1.5million in expenditure [50].

All NHS hospitals in England are required to annually report activity and unit costs using a standard methodology of step-down costing [49]. Costs should be calculated using a full costing methodology, by maximising direct charging and, where this is not possible, using standard methods of apportionment. The costing methodology requires hospital facilities to be grouped together into one of three types "cost pools". Distinction is made between direct costs, indirect costs and overhead costs. This process continues until all of the costs of the

provider are allocated or apportioned to the relevant clinical services. For inpatient and day case activity, the costs are then further disaggregated to HRGs.

Nursing costs are determined as semi-fixed costs in the costs per HRG. It means that costs are fixed for a given level of activity but change in steps, when activity levels exceed or fall below these given levels. Nursing costs are part of the ward cost pool and are apportioned based on the number of bed days.

The costs associated with critical care services are excluded from the composite cost and length of stay for the treatment and procedure (HRG). Any stay in critical care should be extracted from the overall length of stay prior to grouping activity. A separate cost per bed day is produced. Critical Care HRGs are currently in development. It is intended that these HRGs will be based on the number of organs supported, and will supersede current critical care cost and activity data. In an attempt to support such development, level of care data is required for Adult Critical Care services and neonatal intensive care units.

Italy is using DRGs since 1995 for the financing of hospitals [51]. They are using the HCFA /CMS DRGs 19th revision as well on national as on regional level since 2006. Italy has a fixed price financing system based on DRGs but tariffs are different according to the level of care provided and provider settings (teaching hospital, research hospital, general hospital, specialized hospital, etc.). Regional tariffs may vary according to case mix complexity, volume treated, public/private network, and provide extra funds for high specialty providers. Drugs are included in DRG tariffs.

The DRG system finance all inpatient care related to the hospital stay. DRG tariffs include reimbursement of all resources used during the process of care including equipment, personnel (including nursing), drugs, room and board. No weight is applied on DRG tariffs.

Portugal is using HCFA-DRGs version 16 for hospital financing since 1990 [46]. The national Institute of Financial and Information Management (IGIF) is responsible for the financing of the hospitals. There is a standard DRG fixed price. The prices are adjusted according to the hospital structure. DRG financing is used for all acute inpatient care and ambulatory surgery.

The DRG system works with Maryland service weights adapted to the treated patient, length of stay and costs by service. It distributes the total costs for inpatient care per service and DRG. DRG financing includes all health professional costs including nursing care and nurse salaries.

3.4.2 Countries using DRGs without adjusting for nursing care, but with plans take nursing into account

Some countries such as Denmark and USA are not adjusting the DRG cost-weight for nursing care, but are planning to do so. The main reason is that of cost compression, meaning that hospitals that have low nursing intensity patients do better within this reimbursement framework, and hospitals that have high nursing intensity patients tend not to do as well.

Denmark is using the Nord-DRG-DkDRG version for grouping which has been developed with the involvement of medical specialists (the Danish case-mix system is widely accepted by the Danish clinicians) [46]. It consists of two classification systems: the Danish DRG system (DkDRG n=599) and the Danish Ambulatory Grouping System (DAGS) [52]. Basic information is retrieved form the Danish Minimum Basis Dataset for hospitals.

The tariffs attached to the Danish case-mix system reflect the average costs associated with treating the patients in each individual group. The costs include all hospital costs except research, depreciation and capital costs. The average costs of each individual group are calculated by the National Board of Health on the basis of data from a cost database containing cost data from almost all public acute hospitals. Costing of each patient contact is carried out by aggregating the costs of the services consumed by the patient during the contact (i.e. bed days, x-rays, laboratory tests). The model implies that it is possible, systematically to link information concerning the services - and thereby costs consumed - to the individual patient contact. However, the accuracy of the resulting cost per patient may vary between hospitals. In the one extreme, a hospital may be able only to collect information on the number of bed-days, or ambulatory visits per individual patient contact.

In the other extreme, hospitals may be able to link bed-days as well as procedures, examinations and tests to the individual contact. The calculation of cost per unit of service produced: the unit of services defined vary from one cost centre to another, depending of the kind of services produced. For ancillary services, the cost objects may be clinical classifications such as a classification of surgical procedures, classification of radiological procedures etc. For clinical departments (in-patients) the cost objects are bed days, for ambulatories and day case departments the cost unit will be ambulatory visits or day-visits. If the cost unit of the cost centre is not homogenous according to resource use, the costs are allocated to each single unit of service via relative cost weights reflecting the relative costs of producing the different services produced. This is the case for most ancillary clinical services such as x-ray, laboratory tests and surgical procedures. At national level, relative cost weights are calculated for: Surgical procedures, anaesthesia, clinical biochemistry, radiology, pathologic anatomy, physiotherapy and ergo therapy. The cost studies, on which the relative cost weights are based, have been carried out on the initiative of the National Board of Health, often in collaboration with the relevant medical speciality association. There is no national nursing care weighting system. The result is that the cost per unit is calculated simply by dividing total cost by total number of cost units produced.

According to the National Board of Health, on average 71% of the cost per DRG in the cost-weights for 2006, was attributed through bed days. Furthermore, on average 80% of the cost per DAGS in 2006 was attributed through visits. Simply allocating these costs per bed day (length of stay) or per visit may result in a high degree of inaccuracy in the costing of the single patient, as the resources consumed per day or per visit may vary between patients. As a result some studies have been carried out in Denmark to refine nursing cost-weights per DRG [53]. Until now they aren't implemented in the financing system.

The hospital financing system in the **United States** is complicated [54]. The system is mainly a two-tier system whit separate payments for hospitals and physicians. For hospital operating costs, there is no uniform payment system or rates for hospitals in the United States. Medicare pays all hospitals using a common rate-setting methodology, but as Medicaid rates and payment methods are determined by individual states, it leads to different rates to different hospitals.

Beginning in 1983, Medicare implemented the hospital prospective payment system (PPS) [55]. This system changed the basis of Medicare's payments for inpatient hospital care from retrospective costs to a prospective fixed rate per discharge. Cases are categorized by diagnosis-related groups (DRGs). Each DRG is assigned a weight based on its cost relative to the national average cost for all cases. Relative DRG weights reflect the relative rates that Medicare pays for patients' admissions for each DRG case. Basic DRG payments are adjusted (blended rates) according to location (i.e., large urban, other urban, or rural), size, poor population served, services, teaching status and case-mix index.

Next to Medicare, there is Medicaid, targeting low-income families, poor elderly, and the blind and disabled populations. Medicaid now accounts for about 17 percent of total national spending on hospital care. Payment methods vary from state to state, but two methods dominate for inpatient payments: flat fees per DRG or flat per diem payments. Next to Medicare and Medicaid programs, hospitals receive roughly one-third of their net revenues from private health insurers.

Hospital nursing care has traditionally been billed using a fixed daily room and board rate. Nursing care is taken into two cost centers: routine and Intensive Care. There is a separate Nursing Administration cost center as well. This treats nursing care as a fixed cost, e.g. a average cost per patient day for each of the two cost centers.

There is a longstanding interest in the USA to include nursing data in the hospital discharge abstract. The idea of a Nursing Minimum Data Set (NMDS) to help articulate this goal was proposed by Werley in the late 1970s. The main argument for including nursing data in the UHDDS is that medical diagnosis alone does not adequately explain the nursing component of care during the hospital stay. In 1985, the Health Care Financing Administration (HCFA) made available contract funds through the American Nurses' Association (ANA) to begin an investigation of nursing intensity within DRGs [27]. It didn't result in changing the hospital financing scheme until now. Several initiatives have been taken on a local scale. E.g. The New York State Nurses Association (NYSNA) has successfully used nursing intensity to adjust

Medicaid payments in the state by creating a separate Nursing Intensity Weight (NIW) for each DRG grouper [56]. An expert panel is assembled every few years to determine the NIW for each DRG. By legislation, these weights are used to adjust payment to hospitals in New York State (www.NYSNA.org).

In the US, there is now a growing consensus in the nursing community to adjust for nursing intenstiy. The main reason is that there is a cost compression issue, e.g. hospitals that have low nursing intensity patients do better within this reimbursement framework, and hospitals that have high nursing intensity patients tend not to do as well [57]. The Centres for Medicare and Medicaid Services plan to revise the inpatient prospective payment system from a charge based to a cost based formula and to introduce APR-DRGs, which adjust for severity-of-illness starting from 2008 (CMS-1488-P) [58, 59]. The American Nurses Association [60] and American Organization of Nurse Executives support that the proposed CMS-1488-P would adopt adjustment to nursing care as well. However, CMS did not incorporate these into the final rule. Their primary criticism was that there was no nursing specific data in the cost report or corresponding nursing specific revenue codes. The Nursing Minimum Data Set addressed these issues 20 years ago but because there was no implementation of the NMDS (only terminology development), there are no nursing data available [57].

Welton et al. (2006)[61] plead for nursing intensity billing (NIB) in which all nursing costs are directly attributed to patients. NIB for 12 adult medical or surgical units at the Medical University of South Carolina (MUSC) Medical Centre in 2005, resulted in 32,2% increase in charges and a reduction in the variability of nursing cost-to-charge ratios from 0,34 to 0,80 for room and board to 0,33 to 0,45 using the NIB method [62].

3.4.3 Countries using DRGs without adjusting for nursing care now, but with experience in doing so in the past

Some countries such as France and Germany have had experience with adjusting cost-weights for nursing care, but stopped these initiatives.

France is using French DRGs called Groupes Homogènes de Malades (GHM) since 1989 [63]. They introduced a global budget adaptation on DRGs since 1993 and are gradually using them for prospective payment since 2004. The actual system is called T2A (Tarification à l'activité). GHM consists more or less of 800 groups.

National tariffs can be corrected by a geographical factor based on the position of the hospital, status of the organization (profit or non profit hospitals). For non profit hospitals, all professionals' costs are incorporated in the GHM system. For profit hospitals, doctor's fees are not included in the GHM tariffs.

The relative cost-weight of each GHM is based on national costs studies. Yearly, a "National Objective for Hospital Expenditure" is fixed. On bases of the hospital's activity level of the previous year, tariffs per GHM are calculated per sector (Public / Private). The hospital activity is measured in each hospital through the Programme de Médicalisation des Systèmes d'Information (PMSI) [64], which traces all activities done in the organization.

All nursing costs are included in the T2A financing, but there are some special supplements for heavy care activities. Late eighties and begin nineties, France carried out experiments with data collection on the intensity of nursing care [65]. These data haven't been integrated into the T2A financing system.

In **Germany** the Statutory Health Insurance Reform Act introduced in 2000 the system of German Diagnosis Related Groups (G-DRG) [66]. The new reimbursement system will be fully implemented with state-wide base-rates in 2009. The Australian Refined DRG system (AR-DRG system) served as a foundation for the G-DRG system. DRGs are meant to cover medical treatment, nursing care, the provision of pharmaceuticals and therapeutic appliances, as well as board and accommodation. G-DRGs do not cover capital costs. In 2006, there are 914 DRGs with national uniform cost weights, 40 DRGs without national cost weights, and 82 supplementary fees. The 40 DRGs without national cost weights are individually negotiated with each hospital.

The Institute for Hospital Reimbursement (InEK) [67] is responsible for calculating cost weights. The cost weights are calculated using a sampling of data from hospitals participating

in a voluntary data sharing programme. In 2006, a total of 284 hospitals have agreed to participate in the data sharing programme. The participating hospitals must meet certain cost accounting standards. They have to calculate costs per case according to the full cost method using actual costs (Table 8). Based on this matrix costs are allocated to the different DRGs. The drivers used in the cost allocation process differ per cost centre group and cost element group. The "labour cost of the nursing staff" is attributed to the DRGs based on the "PPR-minutes". This system, "Pflege-Personal-Regelung" uses a set of nursing categories which correspond to "expected need of nursing time" was introduced in the German hospitals in 1992. In 1996 it was discarded because the estimated nursing time was "too high" to be accounted for in the financing system.

Table 8: Cost allocation matrix in German hospital reimbursement [64]

Anlage 5		Personal-kosten ärztlicher Dienst	Personal-kosten Pflegedienst	Personal-kosten med.-techn. Dienst/ Funktions-dienst	Sachkosten Arzneimittel		Sachkosten Implantate/ Transplantate	Sachkosten übriger medizinischer Bedarf		Personal- und Sachkosten med. Infrastruktur	Personal- und Sachkosten nicht med. Infrastruktur
		1	2	3	4a	4b[1]	5[1]	6a	6b[1]	7	8
Normalstation	1	Pflegetage	PPR-Minuten[2]	Pflegetage	PPR-Minuten[2]	Ist-Verbrauch Einzelkosten-zuordnung	nicht relevant	PPR-Minuten[2]	Ist-Verbrauch Einzelkosten-zuordnung	Pflegetage	Pflegetage
Intensivstation	2	1. Gewichtete Intensivstunden 2. Intensiv-stunden	1. Gewichtete Intensivstunden 2. Intensiv-stunden	1. Gewichtete Intensivstunden 2. Intensiv-stunden	1. Gewichtete Intensivstunden 2. Intensiv-stunden	Ist-Verbrauch Einzelkosten-zuordnung	Ist-Verbrauch Einzelkosten-zuordnung[3]	1. Gewichtete Intensivstunden 2. Intensiv-stunden	Ist-Verbrauch Einzelkosten-zuordnung	Intensivstunden	Intensivstunden
Dialyse-abteilung	3	1. Gewichtete Dialysen[4] 2. Pflegetage Dialyseleistung	1. Gewichtete Dialysen[4] 2. Pflegetage Dialyseleistung	1. Gewichtete Dialysen[4] 2. Pflegetage Dialyseleistung	1. Gewichtete Dialysen[4] 2. Pflegetage Dialyseleistung	Ist-Verbrauch Einzelkosten-zuordnung	nicht relevant	1. Gewichtete Dialysen[4] 2. Pflegetage Dialyseleistung	Ist-Verbrauch Einzelkosten-zuordnung	1. Gewichtete Dialysen[4] 2. Pflegetage Dialyseleistung	1. Gewichtete Dialysen[4] 2. Pflegetage Dialyseleistung
OP-Bereich	4	Schnitt-Naht-Zeit mit GZF und Rüstzeit[5]	nicht relevant	Schnitt-Naht-Zeit/HLM-Zeit mit GZF[6] und Rüstzeit[5]	Schnitt-Naht-Zeit mit Rüstzeit[5]	Ist-Verbrauch Einzelkosten-zuordnung	Ist-Verbrauch Einzelkosten-zuordnung	Schnitt-Naht-Zeit mit Rüstzeit[5]	Ist-Verbrauch Einzelkosten-zuordnung	Schnitt-Naht-Zeit mit Rüstzeit[5]	Schnitt-Naht-Zeit mit Rüstzeit[5]
Anästhesie	5	Anästhesio-logiezeit[7] und GZF[8]	nicht relevant	Anästhesio-logiezeit[7]	Anästhesio-logiezeit[7]	Ist-Verbrauch Einzelkosten-zuordnung	nicht relevant	Anästhesio-logiezeit[7]	Ist-Verbrauch Einzelkosten-zuordnung	Anästhesio-logiezeit[7]	Anästhesio-logiezeit[7]
Kreißsaal	6	1. Aufenthalts-zeit Patientin im Kreißsaal 2. Anzahl Geburten	nicht relevant	1. Aufenthalts-zeit Patientin im Kreißsaal 2. Anzahl Geburten	1. Aufenthalts-zeit Patientin im Kreißsaal 2. Anzahl Geburten	Ist-Verbrauch Einzelkosten-zuordnung	nicht relevant	1. Aufenthalts-zeit Patientin im Kreißsaal 2. Anzahl Geburten	Ist-Verbrauch Einzelkosten-zuordnung	1. Aufenthalts-zeit Patientin im Kreißsaal 2. Anzahl Geburten	1. Aufenthalts-zeit Patientin im Kreißsaal 2. Anzahl Geburten
Kardiologische Diagnostik/ Therapie	7	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	nicht relevant	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	Ist-Verbrauch Einzelkosten-zuordnung	Ist-Verbrauch Einzelkosten-zuordnung	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	Ist-Verbrauch Einzelkosten-zuordnung	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog
Endoskopische Diagnostik/ Therapie	8	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	nicht relevant	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	Ist-Verbrauch Einzelkosten-zuordnung	Ist-Verbrauch Einzelkosten-zuordnung[3]	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	Ist-Verbrauch Einzelkosten-zuordnung	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog
Radiologie	9	Punkte lt. Leistungs-katalog	nicht relevant	Punkte lt. Leistungs-katalog	Punkte lt. Leistungs-katalog	Ist-Verbrauch Einzelkosten-zuordnung	Ist-Verbrauch Einzelkosten-zuordnung[3]	Punkte lt. Leistungs-katalog	Ist-Verbrauch Einzelkosten-zuordnung	Punkte lt. Leistungs-katalog	Punkte lt. Leistungs-katalog
Laboratorien	10	Punkte lt. Leistungs-katalog	nicht relevant	Punkte lt. Leistungs-katalog	Punkte lt. Leistungs-katalog	Ist-Verbrauch Einzelkosten-zuordnung	nicht relevant	Punkte lt. Leistungs-katalog	Ist-Verbrauch Einzelkosten-zuordnung	Punkte lt. Leistungs-katalog	Punkte lt. Leistungs-katalog
Übrige diagnost. und therapeut. Bereiche	11	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	Ist-Verbrauch Einzelkosten-zuordnung	Ist-Verbrauch Einzelkosten-zuordnung[3]	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	Ist-Verbrauch Einzelkosten-zuordnung	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog	1. Eingriffszeit 2. Punkte lt. Leistungs-katalog
Basiskosten-stelle	12	nicht relevant	nicht relevant	nicht relevant	nicht relevant		nicht relevant	nicht relevant		nicht relevant	Pflegetage

There are still many hospitals that use the PPR-system for internal management. Although the PPR-minutes give an approximation of nursing cost it is criticized because it fails to account for the variability of nursing care. There are different nursing classification systems in use in the German hospitals however without any consequences to the payment system. It is negotiated with the InEK to use systems such as "LEP Leistungserfassung in der Pflege" [5] or the Barthel -index as an alternative in the cost allocation process. It is generally acknowledged that the homogenity of the G-DRGs is not so good with respect to nursing. The assumption is that this homogenity can be increased through better integration of nursing care aspects in the system.

3.4.4 Countries using DRGs with adjusting for nursing care

Some countries, such as Australia, New Zealand and Canada are adjusting explicitly for nursing care.

Australia introduced DRG in early 1990s [5]. **New Zealand** started to use DRG as early as 1988. The use of DRG for payment of health services started in 1995. New Zealand adopted the Victoria (Australia) model. The initial implementation was an adaptation of the HCFA-DRG. Later on, the Australian DRG system was developed up to the actual AR-DRG V5 (Australian Revised DRG version 5) with 665 categories. In Australia, AR-DRG V5 is used as the national standard in describing admitted episodes of care. Although the DRG weighting used to allocate hospital care funding varies from state of state, it represents about 60% in hospital funding allocation. The weights are called "weighted inlier equivalent separations (WIES)" and are annually adjusted (www.health.vic.gov.au/pfg2004/). [68, 69] The National Hospital Cost Data Collection (NHCDC) contains component costs per DRG based on patient-costed and cost-modelled information (http://www.health.gov.au) [70]. The NHCDC enables DRG Cost Weights and average costs for DRGs for acute in-patients to be produced. A hospital can submit either Patient Costed (PC) data or Cost Modelled (CM) data. PC hospitals are able to derive their costs from patient level data. The PC method of costing is also known as a 'bottom up' method of costing because cost aggregates are devised from the individual patient level and summed up to higher levels of aggregation. The information provided by this type of hospital is more detailed and goes down to the individual patient record which includes details like the diagnosis and procedure (clinical costing approach). CM hospitals 'model' their cost centres using pre-determined statistics and weights in order to apportion their costs across product groups and types. This is also known as 'top down' costing because you start with an aggregate cost and apportion it down to the cost centres. [71]

The results of the Collection are published by cost bucket. Cost buckets are the groups against which all costs are mapped. Costs associated with each DRG are reported in the following cost buckets: Ward Medical, Ward Nursing, Non-clinical Salaries, Pathology, Imaging, Allied Health, Pharmacy, Critical Care, Operating Rooms, Emergency Department, Supplies and Ward Overheads, Specialist Procedure Suites, On-costs, Prostheses, Hotel, Depreciation.

The ward nursing cost bucket includes the nursing salaries and wages for each acute stay in general ward areas. Included are nursing salaries and wages reported in clinical service areas. Cost allocation can be done by using service weights or by using patient consumption data. Where data on individual patients' use of resources are available**,** patient consumption data are collected on each patient at the hospital during the Collection period. It is assumed that the nurses routinely score every patient by using a nursing acuity system such as PAIS [72], and allocated nursing costs to each patient in proportion to those data. The data for each patient are the sum of acuity scores from admission to discharge. Other kinds of measures could be used, but they would all have to be correlated with the total nursing cost for each patient.

Where data on individual patients' use of resources are not available, a tool to distribute costs amongst patients using a particular resource is required. For this purpose nursing service weights have been developed, showing the relative use of a particular resource by patients across AR-DRGs. Service weights are derived from studies that measure the typical use of a particular resource by patients in each DRG. Currently there are specific service weights for allied health, audiology, nursing, occupational therapy, pharmacy, physiotherapy, speech therapy, supply, pathology, imaging, CCU, intensive care (adult, paediatric and neonatal), prostheses, operating room, and paediatric and adult nursing. Working on nursing service weights has started by Picone et al. at the Nursing Costing Study in 1993 and has been further refined during the next years [41, 73]. Once service weights have been applied to volume and cost data to derive patient level costs, this data is then used to calculate *cost weights*. A cost weight is the measure of the average cost of a DRG, compared with the average cost across all DRGs. National cost weights are calculated by the NHCDC.

The **Canadian hospital financing** is using a single-source public funding allocated to hospitals via global budgets established by provincial Ministries of Health [74]. The current methodology is comprised of two independent models: the Rate model and the Volume

model. These two models are combined in a multiplicative fashion (*Rate X Volume*) to allow hospital-specific benchmarking of expenditures. The model recognizes both the needs of populations (via the Volume model) and the cost efficiency of providers (via the Rate model).

The outcome of the volume model is the number of expected acute inpatient and day surgery weighted cases for a geographic area. This is based on population demographics and relative needs factors. The key data inputs to the rate model include the Ministry-created Ontario Cost Distribution Methodology (OCDM) methodology which summarizes and allocates annual hospitals costs by various patient categories, and clinical data provided by Canadian Institute for Health Information (CIHI) in the form of a series of databases. The Rate model calculates a hospital-specific expected cost per weighted case for acute, day surgery, and chronic care activity.

These factors, when added to the base rate, gave each hospital an expected cost per equivalent weighted case. The values for all these factors are updated regularly.

For classifying cases, a similar methodology as DRGs are used called Case Mix Groups (or CMG). CMGs are subdivided in level of Complexity comparable to the severity of Illness groups in DRGs. The Expected Length of Stay (ELOS) values are estimates based on the most current patient length of stay information available from the Discharge Abstract Database (DAD) [75]. The ELOS values are adjusted for complexity and age if the adjustments show an improved accuracy of LOS.

The annual calculation of Resource intensity Weights (RIW) requires patient-specific cost data. This is done by CIHI. The Canadian MIS Database (CMDB) contains financial and statistical information from hospitals. The data are collected according to a standardized framework for collecting and reporting financial and statistical data on the day-to-day operations of health service organizations. The framework is known as the *Standards for Management Information Systems in Canadian Health Service Organizations* (MIS Standards) [76]. The MIS Standards are a comprehensive set of standards used to report management information that is ultimately submitted to the CMDB and is related to staffing, costs, workload and provision of services. One element in the MIS-standards that relates to nursing care are Workload measurement systems (WMS) which are time tracking management systems that provides a standardized method of measuring output. The WMS for nursing and most of the therapeutic disciplines moved to a standardized framework for data collection from 1997. There is a WMS for Pharmacy, Clinical Laboratory, Diagnostic Imaging, Respiratory Services, Electro diagnostics, and Non Invasive Cardiology. The MIS Guidelines do not specify a specific methodology for collecting workload data. The framework provided in the guidelines is a reporting framework. Any system which can meet the reporting framework requirements is acceptable for the collection of workload information. Workload must be linked to the functional centre reporting the worked hours of the provider and the patient activity generating the consumption of resources.

Nursing workload reporting has been mandated for Ministry of Health reporting for the fiscal year 97/98. The Nursing Professional Advisory Working Group [77] suggested reviewing the consumption of nursing resources on a broader perspective which would consider not just the collection and reporting of nursing workload but also the factors that influence workload and the role that workload data plays in identifying the contribution of nursing to the provision of patient care. Three nursing workload systems dominate the market (GRASP, NNIS, Medicus) [78]. Only 20% of Ontario hospitals did not have a workload measurement system. Those without workload systems tended to be small hospitals. All of the hospitals with workload systems reported the presence of workload measurement in med/surg inpatient units, about 60% had workload in specialty units and almost 45% had workload in ambulatory care areas.

3.4.5 Countries using DRGs with adjusting for nursing care and calculating the real cost of nursing care on a patient level what would allow to measure the variability of nursing costs within DRGs

Switzerland has introduced DRGs first in 2002 in the cantons of Vaud and Zurich [79]. Swiss DRG should be introduced in the whole countries by 2010. Switzerland is using AP-DRG Version 12. Above the APDRG system, special groups have been created (Swiss Payment

groups or SPG) for special invoices. A suggestion for the Swiss APDRG was to group APDRG and SPG into Activity Analysis groups (GAA), representing 185 groups of care aggregated in 22 major activity groups (PA) based on the chapters of the international diagnostic classification. But this classification is still under study.

The DRG system is based on specific case-mix cost-weights [80]. The weights are based on activity and cost based (all inclusive cost excluding capital costs). An average cost (for university and non-university hospital) is calculated per APDRG.

In determining nursing cost within DRG cost weights, two major nursing workload systems are used: LEP (Leistungserfassung in der Pflege) is used mainly in the German speaking hospitals and PRN (Programme Recherche Nursing) is used mainly in the French speaking hospitals. Still some hospitals are not using any nursing workload system.

A project called Nursing Data [81] aims to develop a national nursing information system for the health sector (as well hospitals, home care, long term care) in all medical specialties and in the four Swiss languages. It should be compatible with all used classifications used in Switzerlland (CIM-10, CH-OP, TarMed, etc.) and should allow international comparisons. This project is realized by the Institut de santé et d'économie (ISE) and resulted in a CH_NMDS. It considers data about the organization, the personnel, the stay, the place of care, the deciding events (diagnostics), the interventions and the link between the event and the interventions. The link between nursing classification and case mix is established via an analytical accounting methodology. For each patient, the allocated nursing time is estimated via the LEP or PRN. This time allocation is converted in francs via the hospital accounting unit where the patient stays. This system allows to estimate a real nursing cost of care per patient. This cost is taken in account in the case cost and the cost weight. A recent study shows that the nursing cost represents around 37% of the total case cost [82].

At the same time, Switzerland is changing this system toward a more German DRG like system which is not favorable at all for nursing costs compared to the actual system.

3.4.6 Countries that are adjusting the hospital financing system for nursing care, but not directly linked to DRGs

Some countries, such as Belgium and Luxemburg are correcting for nursing care but rather independently from DRGs. The case of Belgium has already been discussed in previous chapter. Another example is Luxemburg that isn't using DRGs at all.

Though **Luxembourg** is not using DRG's to finance hospitals, it is interesting to see how nursing activities are measured and used in the hospital budgeting. Luxembourg has 13 acute care hospitals spread throughout the country. Until 1995, hospitals were financed on the basis of a uniform per diem payment, lump sum payments for various surgical operations, and fee-for-service remuneration of physicians. However, a prospective payment system has been in operation since 1995. Hospital budgets are determined individually by negotiation between each hospital's administrative board and the Union of Sickness Funds. The basis for negotiation is the contract between the union of the sick-benefit funds (UCM) and the representative of the Luxembourg hospitals, "Entente des Hopitaux Luxembourgois (EHL)". (www.EHL.org)

Within this contract it is agreed that "…The personnel costs to the UCM are negotiated on the basis of a standard established according to a uniform methodology for all the hospitals. The standards of care-giving personnel are established using the PRN method…". All parties in the Luxembourg system have agreed to use the Canadian PRN system for measuring workload in nursing units.

The PRN [83, 84] calculates the number of staff needed on each nursing ward for the next 24 hours. This method is based on the nursing care plan. It's used in Canadian, Australian and European hospitals (Italy, Spain, France, Switzerland, Portugal and Belgium). Basically, it is a prospective workload study tool. It is based on the assessment of the direct and indirect nursing care needed for the next 24 hours by each patient on the ward. Nursing care and patient needs are grouped into eight categories: Breathing, Eating, Elimination, Hygiene, Movements, Communication, Treatment and Diagnosis.

Each category is further divided into levels of nursing care which contain one or more numbered mutually exclusive factors. The time needed for each of the 249 factors has been defined by studies and expert panels. The sum of each factor's time gives the direct and indirect care time required by a patient during a 24 hour period. The PRN system adds time for the communication about the patient, ward administration, transporting patients, lab samples, or documents to other units ... The addition of all factors calculates the number of full time equivalents needed on this ward for quality care. The PRN method can be used only for inpatient units.

The results of the annual audits are used as negotiating basis to determine the next year staffing. The hospital receives a personnel budget for the inpatient units representing at least 82% of the PRN audit result. Nursing budgets represent from 30 to 45% of the global hospital budget.

3.5 DISCUSSION

From the literature overview it is clear that there are many different systems in use to take care of nursing care in the hospital financing system. It is clear that countries are moving back and forward. Germany was adjusting its hospital financing systems for nursing care until 1996, but declined afterwards. The USA is making plans to introduce a nursing care into the financing system again after several trials during the last 20 years.

The main reason for adjusting the financing system for nursing care is that of cost compression. The nursing factor is so large (20 – 30% of all costs), that using the average nursing cost gives a real bias by overestimating lower cost patient groups and underestimating higher cost patient groups. The impact of this adjustment is debatable. Cromwell & Price (1988)[85] showed in 1988 that the impact of adjusting DRG cost-weights with nursing intensity is limited. Although the impact on individual DRGs is great, the impact on the level of the hospital is limited. 95% of all hospital budget would not change more than 1% in either direction (+1%, -1%). The main criticism on this analysis is that the analysis is cross-sectional on one moment of time. It would be interesting to evaluate what would happen with changing case-mix and length-of-stay.

The main limitation for adjusting for nursing care is certainly the availability of data. Most countries struggle with the availability of uniform nursing data. Most clinical costing methods make use of the available nursing workload systems within hospitals. In most countries different nursing workload systems are accepted, but require one reporting standard. When these data are not available, nursing services weights are an alternative. In many countries, these service weights exist for many different cost centres. The Australian and New York experience show that it is also possible to develop these nursing service weights for nursing care.

Most of the countries which are adjusting DRG-cost weights for nursing care, stop at the level of an average cost weight per DRG. Some experiences, such as the Nursing Data experience in Switzerland or the Nursing Intensity Billing experience in the US, show that it possible to link DRGs and nursing data. This would allow investigating further the variability in nursing care within DRGs. The limited explained variance (15% to 20%) show that more research is needed for further investigation of this relationship and that averaging nursing care per DRG is probably a too crude measure.

4 EVIDENCE BASED CHARACTER OF NURSING INTERVENTIONS AS DEFINED IN THE B-NMDS-II

4.1 INTRODUCTION

The process of revising the Belgian Nursing Minimum Dataset (B-NMDS) started in 2000. Based on a literature review and secondary data analysis, the Nursing Interventions Classification (NIC) was selected as a framework for the revision of the original B-NMDS [86]. After discussions in panels of clinical experts (*N* = 101) and extensive testing in 232 nursing wards in 66 Belgian hospitals in December 2003, February 2004, March 2004 and March 2005 (N=117395 in-patient days), 78 nursing interventions were selected in the final B-NMDS, second version [1].

Evidence-based practice is rapidly growing in healthcare. But in nursing practice the use of research findings is not widely spread. Bostrom & Suter (1993)[87] found that only 21% of 1200 practicing nurses had implemented evidence from a research study into their practice during the previous 6 months. Findings of a survey of nurses in western Canada illustrate that nurses use a broad range of practice knowledge, much of which is experientially based rather than research-based [88].

The reason for incorporation an EBN evaluation within the framework of the nursing minimum dataset is that there is a high probability that many of the nursing interventions are not always based on hard evidence. In a qualitative study of Thompson et al (2000)[89] in the UK, it was found that most decisions that nurses made or reported were focused on areas such as: dressings, pressure sore monitoring/ prevention/ use of devices, checking/monitoring observations/ fluid intake and output, patient hygiene, patient mobility, patient positioning, infection control, nutrition, iv/oral fluids, timing of preoperative medication, patient drug compliance, referral to colleague/ senior nurse/ doctors/ clinical nurse specialists/ therapists/ pharmacists, referral to relatives/involvement of relatives especially around discharge, interpreting results, mainly of blood tests/exercise tolerance tests and decision to document care given. Many of these areas are common for the NMDS. Their conclusion is that "not all of them merit a full scientific experiment but neither do they all merit intuitive guesswork". An introduction of more evidence based decision making would not only impact today's nursing practice but also nurse staffing levels and reimbursement (in plus as well as in minus).

4.2 RESEARCH QUESTION

However, any step towards an 'appropriated care' financing system for hospital nursing care is bound by the availability of evidence based clinical recommendations concerning hospital nursing care. Without sufficient evidence generation in the field of nursing, a further application is premature. Therefore this chapter is aimed at the following research question:

What is the evidence based character of the nursing interventions as defined in the B-NMDS-II?

4.3 METHODOLOGY

4.3.1 Selection of NMDS - items

Out of the list of 76 nursing interventions from the NMDS-II, a limited set was selected. The selection was based on 4 criteria:

- The frequency of which these interventions occur in the Belgian hospitals. All interventions were ranked according to their prevalence (x / 1000 inpatient days) for surgical and internal medicine departments.

- The variability of care: in some hospitals/ nursing units specific nursing care interventions occur with a much higher prevalence than in some other hospitals/nursing units. The hypothesis is that the variability of care is related to differences in practice or differences in coding. The coefficient of variability has been used as measure for variability. All interventions were ranked according to this ratio.
- Relationship with nurse staffing: some nursing interventions have a higher requirement of nurse staffing than other nursing interventions because they require more time to perform, have a high daily frequency or have to be executed simultaneously by more than one nurse to assist the patient. Relationship with nurse staffing was evaluated on a five point Likert scale ranging from 1 (absolutely no relationship with nurse staffing needs) to 5 (strong relationship with nurse staffing needs). The evaluation was done by an expert panel.
- Awareness of existing evidence for the selected nursing intervention. This was evaluated on a five point Likert scale ranging from 1 (great lack of scientific knowledge for evaluating evidence for or against) to 5 (ample scientific knowledge for evaluating evidence for or against). The evaluation was done by an expert panel.

The evaluation of criterion 1 and 2 was performed based on the data collected during the pilot phase of B-NMDS-II. The evaluation of criterion 3 and 4 was performed by using an expert panel. The panel consisted of seven nurses with expertise in the field of evidence-based nursing (EBN).

Based on the four criteria, fifteen different nursing interventions have been selected. In the research team the selection was discussed. A final selection of nine different nursing interventions were selected for EBN review: B230; B300; K300; V600; F500; E100; V100; V200; V700. This selection is indicated in grey in Table 9.

Table 9: Selection of nursing interventions for EBN review*

Nursing interventions	Frequency (per 1000 inpatient days)	Variability (Coefficient of variability and Standard Deviation)		Average staffing need (1-5)	Awareness of existing evidence (1-5)
Urinary incontinence care (B230)	914	14	196	4,4	4,4
Bladder catheterization (B300)	17	107	5,7	3,8	4,8
TPN administration (D400)	33	147	17,9	3,8	4,4
Artificial ventilation (K300)	16	316	14,9	4	4,6
Administration of blood (N100)	15	92	9,6	3,8	4,5
Hygienic care (F110, F120, F200)	807; 23; 38	19; 181; 345	175,4; 10,5; 36	3,8; 4 ; 3	1,8 ; 1,8 ; 2,6
Wound care (L100,L200,L300)	0; 241; 106	0; 67; 65	0; 28,6; 9,6	2 ; 2,5 ; 2,8	5; 5; 4,4
Patient education (S100, S200)	144; 82	134; 109	55,7; 29,2	4,4 ; 4	3,8 ; 4,4
Isolation care (V600)	33	143	14,9	3,8	4,6
Special mouth care (F500)	119	170	55,8	3,4	3,6
Symptoms management pain (E100)	359	93	122,9	3,2	4,6
Glycemic regulation (G300)	146	66	39,7	3,2	3,8
Ulcer pressure prevention using dynamic systems (V100)	119	113	21,8	2,8	4,8
Ulcer pressure prevention by repositioning (V200)	61	116	40,3	3,8	4,8
Protective measures with des orientation (V700)	58	159	20,3	3,8	4,6

*(Grey = selected for EBN review)

4.3.2 Literature review EBN

Based on KCE E.B. guidance and consultation of CEBAM experts the framework for the literature review was determined.

4.3.2.1 Formulating the research question

In a first step, per nursing intervention, a PICO was formulated.

- P = Patient(group) characteristics, population (demographic: age, gender; disease: pathology, DRG, comorbidity; setting; treatment stage)
- I = Intervention
- C = Comparative intervention/golden standard
- O = Outcome

4.3.2.2 Search strategy

The search strategy focused on three forms of synthesized literature within a wide spectrum of databases, portals and other sources:

- Guidelines
- Systematic reviews
- Specific EB(N) sources

Fourteen guideline sources, three systematic review sources and two EBN specific sources were searched systematically. Sources are presented in Appendix 2. Based on an iterative approach, original studies were searched additionally if the sources above yielded an insufficient number of publications. This additional search was performed using the databases of Centre for Research and Dissemination (CRD), Cinahl and Medline.

4.3.2.3 PICO refinement and Screening

The in- and exclusion criteria are defined specifically for each NMDS intervention. The screening was divided by intervention over 6 reviewers. If the compliance with in- and exclusion criteria was doubtful, an independent reviewer was consulted. A small sample of publications was independently reviewed by two persons to ensure a similar approach by all reviewers.

4.3.2.4 Quality Assessment

For "trusted" sources such as NICE, SIGN, CBO, JBI, WVVH, NCCHTA, Duodecim, CEBAM-LIBRARY, CDSR, DARE, Clinical Evidence, Evidence based nursing, and ICSI no explicit quality assessment was performed. For the results that are retrieved in other sources of literature (see above), a systematic quality assessment was applied. The process of quality assessment is presented in Appendix 3.

4.3.2.5 Data extraction and level of evidence

For each publication following data are retained:

- Source of the document
- Reference
- Type of study : guideline/systematic review/other EB(N) source
- A list of indications and contra indications for the use of the intervention. This mostly corresponds with 'P' characteristics such as:
    - Age and gender of patient group
    - Diagnose, pathology, DRG, co morbidity if described

- Setting (hospital, type of ward, primary care,…)
- Treatment stage if described

- Intervention: All elements concerning the NMDSII intervention itself, mostly in 'I' terms
  - Specific description of and recommendations about the intervention components
  - A list of indicated context interventions
  - A list of contra indicated context interventions
- Comparative intervention used and Outcome:
  - Aimed comparisons and outcome measures were extracted if available. Most publications applied a grading system of levels of evidence, which implied a positive outcome compared to the use of no or alternative interventions. The stated levels in the publications are translated into three general levels of evidence (Box 1).

There is no one best approach in grading evidence [90]. Formally grading study quality and rating overall strength of evidence can however produce reasonable levels of confidence about the science base of study findings. A system similar to the GRADE system was adopted [91, 92, 93, 94], with the exception that grade D evidence (any other evidence, very low quality) has been excluded or incorporated into the C level. A comprehensive quality assessment already took place (see section 4.3.2.4.). Therefore, the study design was the primary driver for further evidence grading. A randomized clinical trial is rated as the highest level of evidence, because bias is controlled for in the best way. Case series are at the other end of the spectrum. The hierarchy within the range of study designs can be consulted in Brighton et al (2003)[95], Busse & Heetveld (2006)[96], Concato et al (2000)[97], Petrisor et al (2006)[93], and standard scientific research textbooks. Simplicity, transparency, explicitness of methodology and consistency are the most important criteria for a grading system [98]. To our belief and based on practical experience, the hierarchy in study design used is clear cut in its process and results.

Box 1: General system of evidence grading

- **A = Good level of evidence: meta – analyses, systematic reviews and rct's**
- **B = Fair level of evidence: ct's, cohort, case – control, before – after studies**
- **C = Weak level of evidence: observational, case studies, expert opinion, consensus, common practice**

4.4 RESULTS

4.4.1 Ulcer pressure prevention (NMDSII items V100 and V200)

The following refined PICO matches NMDSII item specifications:

- **Patient characteristics/Population:**
- Patients that are at risk of developing pressure ulcers.
  - Adults
  - In a general hospital environment
  - Applicable within the Belgian health care context

- **Intervention characteristics:**
- Prevention of pressure ulcers.
    - No treatment
    - Including risk assessment by measurement or clinical judgment, with the use of dynamical materials (beds and mattresses) or alternating positioning. Or interventions that are necessary/avoidable in combination with these NMDSII interventions.
    - Nursing staff contributes to or executes the intervention.
- Comparison intervention:

A specific alternative intervention or no intervention

- **Outcome:**

An evidence based indication of a positive, negative or an absent effect of intervention or a sufficient level of evidence stated to justify an intervention or no intervention

Based on this PICO the following evidence based recommendations can be formulated:

Evidence level A

- Patients who are bed and chair bound, have an impaired ability to reposition themselves, are immobile or have a limited activity level, are at risk for developing pressure ulcers. Further interventions should be targeted to these patients to prevent pressure ulcers.
- For patients who comply with the primary risk factors additional secondary factors have to be taken into account in addition to the physical and functional status. An altered level of consciousness, incontinence or moisture and dietary intake or nutritional status has a clinically significant impact on pressure ulcer development. These issues should be assessed and addressed appropriately. Patients with a combination of primary and secondary risk factors are highly at risk for developing pressure ulcers. In this it's not totally clear what constitutes the factor dietary intake or nutritional status.
- Patients at risk for pressure ulcer development should be repositioned adequately and regularly as a part of primary and secondary prevention. A frequency of every four hours in combination with a special mattress is recommended. If such a mattress is unavailable, a patient should be repositioned every two hours.
- Patients at risk for pressure ulcer development should be placed on a special mattress.
- For high risk patients dynamic systems (large cell alternating, low air loss, air fluidized) are preferred over static systems. In the rare situation that repositioning is impossible due to clinical objections, the use of a dynamic system is a prerequisite.
- The use of low air loss hydrotherapy beds should be avoided.
- Some higher level studies show a significant effect of high protein and calorie supplements in specific populations (e.g. older patients).
- The use of zinc and ascorbic acid supplements is not supported.
- Finally, there is spares high level evidence which indicates that a timely transfer to enteral or parenteral feeding could be beneficiary for pressure ulcer prevention.
- There is strong evidence available that supports the effect of education of the patient and his family in preventing pressure ulcers. The content includes topics such as etiology, risk factors, risk assessment, skin assessment, skin care and positioning.

- High level studies denounce the use of chlorinated solutions.

Evidence level B

- Limited evidence is available that supports a broadening of risk factors to include all illnesses that are severe chronic or terminal in nature, or generally imply an acute immobility. So a translation in specific medical diagnostic terms alone should be avoided. A patient by patient approach is warranted, with a focus on functional ability. This should be assessed in a scientific manner.
- Protein and calorie malnutrition and dehydratation are supported as constituents. Moisture refers to a balance between avoiding a too dry skin and an excessive exposure to moisture caused by incontinence, perspiration or wound drainage fluids.
- The impact of the following characteristics is supported: age (> 75 years), a history of ulcer development, friction and shearing, an impaired sensory perception (loss of feeling in certain parts of the body, comatose status or lack of sensation). The same holds for the diagnoses diabetes mellitus, polyneuropathy or vascular disease, fever, paralyses and the use of medication such as an extensive antibiotics treatment. The skin should be inspected for non blanching erythema which shows a high predictive value for pressure ulcer development.
- Some recommendations are given concerning the frequency of assessment, such as 'within two hours after admission' and 'with a frequency of every 72 hours'.
- In executing the repositioning the avoidance of extensive massage and the relief of heels are supported.
- There is also some indication that cubes, hollow fiber mattresses and water mattresses should be avoided.
- Massage over bony prominences should be avoided, because this can lead to deep tissue trauma.

Evidence level C

- The following factors are supported by weak evidence only: non cognitive psychological status (mood, motivation, aptitude), pain status, social factors, the existence of decubitus, elements concerning skin status next to non blanching erythema, diagnostic factors such as pregnancy, obesity, systematic signs of infection, spinal cord injuries, Guillain Barré, Multiple Sclerosis, anemia, myocardial infarction, stroke, multiple trauma, musculoskeletal disorders, fractures, gastro intestinal bleeding, renal disease, cancer, COPD, CHF, dementia, preterm neonates. There isn't any fair evidence for the following interventions as risk factors either: the use of medication such as immunosuppresiva and anaesthesics, surgical interventions, a long duration of surgery, the use of a prosthesis, body brace or plaster cast.
- No specific contra indications are stated in scientific literature. Unmotivated treatment and costs should however be avoided for the benefit of the patient and the society.
- Risk assessment should take place at intake and at changing patient status concerning one of the primary and/or secondary risk factors.
- Use of the Norton or Braden scale is recommended as a means for systematic evaluation, although the test characteristics should be improved in future research. There is more evidence available supporting the Braden scale, compared to the Norton scale.
- The combination with clinical judgment and with systematic skin inspection is recommended. These activities can be integrated in other regular care interventions, but shouldn't be overlooked. Education for self inspection is an interesting avenue suggested by some authors.

- Proper positioning, transferring and turning techniques with the minimization of friction and shearing are considered to be self evident by panel consensus. Regular control of the patient's position, control of the bed surface, proper positioning of the head of the bed and the use of lifting devices are weakly supported as preventive for pressure ulcer development. The same applies to teaching patients to move their weight regularly and the use of frequent small position changes.
- Considerations concerning the sitting position of the patient are theoretically grounded, but also show a weak body of evidence. The duration of sitting and the frequency of repositioning during sitting aren't consistent in scientific literature. However, the theoretical grounding implies more stringent demands than during other patient positions in terms of duration and frequency.
- The specific choice between different special systems remains unclear.
- There is weak evidence available that indicates the need for the use of special materials such as an air or foam cushion during sitting. In some studies no significant difference could be found concerning this approach. Hollow fiber, water, gel, sheepskin and rings should be avoided.
- The use of a special mattress doesn't exclude other preventive measures, such as repositioning and other interventions. Special mattresses often imply a higher complexity of care and a higher need for training and education.
- There is weak evidence available which promotes the adaptation of hygienic care by individualization of frequency of hygienic care in function of specific needs (e.g. at time of soiling) and by adjusting the way in which hygienic care is delivered (clean gently, no hot water). Herein the balancing hydration of the skin is a point of interest: the use of moisturizers, creams or lubricants when the skin is too dry and the separation, dry keeping and assessment for candidiasis of skin folds. In this specific care products can be used, such as talc and/or mycolytic sprays.
- Experts advise the minimization of skin moisture due to incontinence in general elimination care and the specific care for risk areas on the skin. The former by using barriers such as underpads and frequent changing. The latter by applying topical agents, protective dressings, hydrocolloids or protective films. A broad package of incontinence treatment to ameliorate the problem is also advised.
- Finally, based on expert opinion, the contact of the skin with wound drainage and other bodily fluids should be avoided as much as possible. Appropriate wound dressing and frequent changing is indicated.
- A good nutritional balance should be garded in all patients, including a sufficient fluid intake. Serum albumin testing is suggested by some authors to assess the patient's status.
- Weak evidence is available about the effect of rehabilitation by motion exercises, ambulation, etc. Based on theoretical grounding and expert consensus however, this is considered an important point of interest. Stimulation of activity participation complements this and reinforces the general coping ability of the patient.
- The avoidance or elimination of pressure due to care aids such as oxygen masks, tubing, catheters, cervical collars, casts and restraints is recommended.
- The whole set of assessment and pressure ulcer prevention interventions should be implemented on a 24 hours of continuity basis.
- Programmed education and occasional education during other care interventions is recommended.
- Some experts suggest pressure ulcer prevention as a point of consideration in discharge planning.
- Donut type devices, ring cushions, ice, sheepskin and preventive bandages over healthy skin should be avoided.

- Prevention should be focused on patient centered outcomes such as the incidence of pressure ulcers (e.g. signs of skin breakdown, ulcer recurrence). Surrogate outcomes such as interface pressure aren't considered to be a valid indicator of the patient's needs, health evolution and effects of care. Adverse events should also be monitored. Attention for patient comfort and quality of life (e.g. acceptability of devices) is required.

4.4.2 The use of protective devices (NMDSII item V700)

The following refined PICO matches NMDSII item specifications:

- **Patient characteristics/Population:**
- Patients, disoriented in time and/or space, with a protective need concerning themselves or others.
    - Adults or elder patients
    - Disorientation as measured with a validated instrument
    - In a general hospital environment
    - Applicable within the Belgian health care context
- **Intervention characteristics:**
- A mobility restriction
    - Seclusion within a room or ward, or the application of a mobility restrictive physical device as part of the intervention
    - No fracture fixation (e.g. splinting and casting) or chemical restraints (analgesia, tranquilizing agents, neuromuscular nondepolarizing agents) as a single focus
    - Nursing staff contributes to or executes the intervention
- Comparison intervention:

No mobility restriction

- **Outcome:**

An evidence based indication of a positive, negative or an absent effect of the intervention or a sufficient level of evidence stated to justify the intervention or no intervention

Based on this PICO the following evidence based recommendations can be formulated:

Evidence level A

- The prevention of falling isn't a valid reason for the use of restrictions. Strong evidence suggest the opposite effect: falling due to restraints, with more serious injuries.
- Incontinence can be treated effectively to ameliorate the cognitive status of the patient.

Evidence level B

- Physical restraint use can lead to falls, serious injury and deaths. The use of these measures should be minimized as much as possible by applying alternatives. Physical protective devices should be reserved as a last resort, after everything else failed. The implementation of spatial boundaries by securing a patient's room or the ward as a whole isn't restricted in this way.
- The following reasons aren't considered valid for physical or spatial restrictions: interference with therapy (e.g. tampering with devices, maintaining patient's position, enabling activities of daily living), provide quiet time for the patient,

disruptive behavior (e.g. bothering others in a non aggressive way, taking things, impulsive behavior).

- The following predisposing factors can be identified: cognitive impairment (e.g. in orientation, attention, speech, judgment, caused by dementia or delirium), the use of medication and related reactions with psychoactive effects (e.g. polypharmacia).
- An active screening for risk patients is beneficiary. Careful observation to identify the behavior is indicated.
- Behavior modification (e.g. stimulus control reinforcement) and light therapy as a part of psychological interventions generate a positive effect.
- There is fair evidence for the effectiveness of activity programs including occupational and recreational activities (e.g. outdoor walks).
- Another focus is the cognitive status of the patient. The effectiveness of Reality Orientation Training (e.g. environmental cues, orientation points) is supported.
- Fair evidence supports the improvement of the functional status by gradual ADL stimulation and physical activities such as light exercise and walking.

Evidence level C

- The following circumstances may allow for the use of physical restraints: the potential for suicide, the potential for violence, the protection of life sustaining treatment, the prevention of physical exhaustion by extreme psychomotor agitation. Wandering, elopement behavior and agitated, restless behavior can be indications for non physical, spatial restrictions.
- It's self evident that reasons as enabling efficient work schedules, nursing comfort, punishment for non compliance and reducing legal liability are invalid.
- Experts add the following as predisposing factors: age in terms of the frail elderly, living in a resident setting, impaired functional status (e.g. activities of daily living, mobility), psychological status (e.g. depression, anxiety, substance abuse), social status (e.g. communication ability, attitude of caregivers), sensory deficits (e.g. seeing and hearing impairment).
- The behavior which leads to protective measures is often precipitated by certain factors, mostly in terms of a source of internal or external stress. Experts mention hunger, pain, need to toilet, fatigue and excess demand or activity requests as examples. An unfamiliar environment or even the use of restraints in itself can give way to the negative behavior. The cause can also be of a physiological nature (e.g. infection, nutrition deficit, electrolyt disturbance, acute physical illness, constipation, incontinence, urinary retention, catheter use, pressure ulcers, changing medication). Often more than one factor underpins the behavior.
- Validated tests such as a behavioral rating scale can be used as a screening aid. Diagnostic interventions comprising an anamnesis and hetero anamnesis, a physical examination, lab tests and other investigations are indicated to detect the underlying needs.
- Close monitoring and information gathering takes place during all patient contacts. Some authors suggest a thorough assessment on a daily basis.
- Patient sensitivity in terms of knowing the person is important in detecting root causes. Four elements can contribute to this empathy: reviewing the patient's history related to past events and coping; discerning triggers and behavior patterns in terms of place, time, persons, antecedents and situations; directly asking the patient for his or her needs; involving the family in the interpretation of behavior and triggers.
- Reversible causes are treated and avoidable triggers are minimized.

- Experts highlight the significance of psychological and social interventions such as: sufficient interactions and conversations, adapted communication techniques (e.g. naming, speed, simplicity), active listening and empathy, emotional support, individualized social contact and activities, provision of appropriate outlets, a 24 hours attitude of support.
- The multi intervention program should also include environmental interventions. Experts suggest the balancing of stimuli according to the patient's needs (e.g. noise, lighting, private space, visitors). There is weak evidence for the use of distraction and diversion by music, pet therapy, etc. The use of white noise and natural elements is experimental with few reports of positive effects up to now. One author suggests night time activities for patients with a disturbed sleep pattern.
- Wandering should be allowed within specific areas. The patient's room should be close to the nursing station to increase patient visibility.
- The following interventions are also beneficiary: simulating a home or family like setting (e.g. familiar people, things and activities) with caregiver consistency in assignment and accommodating the care schedule (sleeping, eating, bathing) to the usual routine of the patient, reminiscence about old times, validation therapy, family participation, the implementation of routine and structure.
- Next to state of the art interprofessional treatment to tackle the physiological problems which trigger behavior, experts mention the use of some more specific leverages: medication assessment and management (e.g. discontinuation), the use of sensory assistance devices (e.g. referral to specialist), maintaining a healthy fluid and nutritional balance, sleep therapy, prevention of pressure ulcers, pain assessment and management, appropriate toileting and incontinence treatment (e.g. voiding and cleansing promptly, individualized rounds, scheduled toileting).
- Furthermore, education of patient and family, responding quickly when care is needed and changing bothersome treatments (e.g. lines and tubes) as soon as possible, is recommended.
- Some authors highlight a higher time requirement for nursing care aimed at risk patients, e.g. caused by the need for increasing relaxation during care activities and an unhurried approach towards the patient.
- A physical restraint must be applied in a standardized manner, using special order forms to start with. An 'as necessary' approach should be avoided.
- The patient and family are notified promptly, explaining the need for this course of action. Informed consent must be obtained within 24 hours.
- The least invasive option, the minimal level of restraint is used.
- A restraint is positioned properly. It is always considered as a temporary measure. The order is limited in time, e.g. 24 hours. Afterwards reassessment is indicated. Caregivers, including nurses, need to take any actions to remove the restraint as soon as possible.
- Whenever direct supervision is available from family or caregivers, the restraint can be removed.
- The use of restraints demands continuous monitoring with a high frequency. Authors suggest a time frame that varies from every 15 minutes up to every four hours in increasing nurding rounds. Clinical condition, orientation, correct placement of restraints, circulation, motion and sensation are all monitored. At those moments the patient should also be released for a short interval of time for reasons of security, comfort and stimulation of mobility.
- The restrained patient should be left alone at least as possible, unless overstimulation adds to the problem. There are authors who suggest the need for a continuous one to one supervision by staff, family, friends or volunteers.

- Patient centered outcomes are the priority. The safety of the patient and others, measures of behavior change (e.g. frequency, severity), patient anxiety or distress and quality of life are examples. The acceptability of the intervention, the experience of being restrained or having a relative restrained adds to these when it comes to the intervention itself. Adverse events should be monitored. On a more aggregate level other indicators can be implemented such as the injury rate, the incidence in restraint use and the prevalence of geriatric syndromes (e.g. delirium).

4.4.3 Symptoms management pain (NMDSII item E100)

The following refined PICO matches NMDSII item specifications:

- **Patient characteristics/Population:**
- Patients that are at risk of developing pain symptoms or already suffer from pain
    - Adults
    - In a general hospital environment
    - Applicable within the Belgian health care context
- **Intervention characteristics:**
- Systematic follow up of pain
    - No treatment of pain as a single focus, without assessment
    - Including measurement using an instrument, peer reviewed in a scientific journal. Or interventions that are necessary/avoidable in combination with this NMDSII intervention. Pain isn't just used as an outcome measure to assess another intervention (e.g. medication studies).
    - Nursing staff contributes to or executes the intervention.
- **Comparison intervention:**

An unsystematic follow up of pain or no follow up of pain

- **Outcome:**
- An evidence based indication of a positive, negative or an absent effect of intervention or a sufficient level of evidence stated to justify an intervention or no intervention

Based on this PICO the following evidence based recommendations can be formulated:

Evidence level A

- Self report by the patient is the primary source of information, combined with physiological (e.g. sweating, cardiovascular changes) and behavioural (e.g. facial expression, restlessness) indicators.
- Strong evidence supports the positive effects of education on the pain outcome of the patient.

Evidence level B

- A specific and thorough assessment should follow after an initial positive screening, prior to treatment. Studies indicate a need for a broad attention towards the physical, functional and psychosocial dimension. This includes: information about the meaning of pain, preferences, expectations, held beliefs and myths about pain, its management and outcome by the patient and family, checking for the presence of depression or anxiety, the type of pain (somatic, visceral, neuropathic), the distress caused by the pain (i.e. the pain tolerance), aggravating and relieving factors.

- Fair evidence studies indicate the need for a prompt reassessment when intensity of pain increases and according to the distress for the patient.
- The following parameters should be reassessed: pain location (e.g. body map diagram), pain quality, pain intensity (at rest and during activity/movement), radiation, timing (occasional, intermittent or constant), duration, effect on function, ADL, activity, sleep and mood, medication usage, level of sedation.
- If the pain persists or is too severe a more comprehensive assessment should take place, making use of the broad array of diagnostic interventions.
- The use of a validated pain measurement instrument has a positive effect on the patient's outcome. There are many standardized tools with established validity available. the use of one dimensional measures of pain intensity or relief such as the VAS, NRS, VRS, VDS or VNRS is supported. The use of multidimensional measures is also substantiated. The Memorial pain assessment card, the McGill Pain Questionnaire and the Wisconsin Brief Pain Inventory are examples of the latter.
- Patient and family input and participation about goals and options in pain management has proven to be beneficiary.

Evidence level C

- Postoperative treatment and cancer treatment are examples of conditions which warrant a secure pain assessment and management.
- The age, mental status, language and cultural background of the patient influence the way in which pain assessment and management can be delivered.
- No specific contra indications for pain assessment and management are described in literature.
- Continuous screening for the presence of pain in patients at risk is indicated.
- Experts suggest a screening at admission and routinely during the whole stay, at least once a day.
- For a comprehensive diagnosis and planning the following characteristics should also be assessed: the pain location, the pain quality (e.g. 'stabbing', 'burning'), coping responses (including tests for plasma cortisol as indicator of the stress response), effects on functional abilities such as coughing, walking and activities of daily living. The presence of dysaesthesias, hyperalgesia, allodynia or hypoaesthesia can be detected. The pain history of the patient, the associated disability and previous treatments and their effectiveness are also of interest. Finally, the education need has to be evaluated according to the preferences and responses of the patient.
- Individual goals for pain treatment should be set from the patient's perspective (e.g. 33 to 50% decrease in pain intensity, taking into account patient satisfaction and the ability to resume reasonable activities).
- In the following situations reassessment is indicated: after implementing a new diagnostic or therapeutic procedure (e.g. initial 24 hours postoperative at least every 2 to 4 hours, during upward titration every 15 minutes), according to the duration of pain, at instances of unexpected pain and when a pain management intervention reached peak effect: after 60 minutes for oral immediate release medication, after 4 hours for sustained release medication or transdermal patch, after 15 to 30 minutes for parenteral medication and after 30 minutes for non pharmacological pain interventions. Furthermore it is suggested to routinely monitor pain together with other vital signs. If pain is out of control, there should be a daily change in treatment until the pain is controlled.
- Experts also suggest the following: regional autonomic features (e.g. colour, temperature), intensity at its worst in past 24 hours, intensity during last week,

extent of pain relief, pattern of pain response, provoking or precipitating factors, barriers, quality of life, side effects of interventions, respiratory rate.

- Pain duration longer than six weeks or longer than the anticipated healing time should trigger an evaluation of the presence of chronic pain.
- Face and behavioural scales are recommended as alternative assessment scales.
- A positive pain assessment promptly should lead to renewal of treatment planning and execution. All caregivers should advocate on the behalf of the patient to ensure a continuing needs based treatment. This isn't only aimed at pharmaceutical treatment, but also on interventions such as repositioning, cuteaneous stimulation, massage, deep breathing, distraction, relaxation, and music, which can be applied by nursing personnel immediately.
- An appropriate nursing patient interaction concerning pain and its management is also considered to be important.
- A combination of planned sessions and occasional education during care interventions is suggested. For patients with an avoidant coping style (i.e. having an external locus of control) education should be contained. Education should start before the onset of pain, together with other pain preventive interventions.
- By some authors symptoms management pain is named as a special topic to consider during discharge planning.
- Within the avenue of a synergistic approach, the use of a standardized program with multiple interventions is advised.
- A focus on multiple outcomes is required. Pain intensity or relief, distress and functional ability (i.e. pain associated morbidity and interference) are the main outcomes. Satisfaction and quality of life complements this. The healing rate or recovery time is another measure. Adverse drug reactions and adverse events should be monitored. On a more aggregate level the following indicators can be useful: analgesia prescription and usage, wait time, length of stay, rate of referrals to pain specialists, readmission frequency.

4.4.4 Isolation care (NMDSII item V600)

The following refined PICO matches NMDSII item specifications:

- **Patient characteristics/Population:**
- Patients that have a MRSA infection.
    - In a general hospital environment
    - Applicable within the Belgian health care context
- **Intervention characteristics:**
- Patient-isolation measures are taken.
    - No treatment
    - Nursing staff contributes to or executes the intervention.
- **Comparison intervention:**

No measures or insufficient measures are taken.

- **Outcome:**
- Less contamination or dissemination

Based on this PICO the following evidence based recommendations can be formulated:

Evidence level A

- Hands must be decontaminated to avoid cross infection, preferably with an alcohol-based hand rub, between caring for different patients or between different care activities for the same patient. If the hands are visibly dirty or contaminated with proteinaceous material or are visibly soiled with blood or other body fluids, they should be washed with either a non-antimicrobial soap and water or an antimicrobial soap and water.

Evidence level B

- Early identification of MRSA reservoirs is essential to the implementation of focused strategies to eradicate the vectors. Periodic (e.g. weekly, according to local prevalence of MRSA) surveillance cultures are indicated for patients remaining in the hospital at high risk for carriage of MRSA, because of ward location, antibiotic therapy, underlying disease, duration of stay, or all four.
- Interventions that include isolation of MRSA infected patients in a separate room or ward, can reduce MRSA transmission. There is however no evidence that isolation of the infected site, if possible, is more or less effective than strict patient isolation.
- There is no evidence that demonstrates the effectiveness of protective isolation in reducing nosocomial MRSA acquisition.
- Gowns should always be worn as part of contact precautions for all patient and environmental contact with patients known to be colonized with MRSA.
- Hands must be decontaminated immediately after any activity or contact that could potentially result in hands becoming contaminated, and after removing gloves.

Evidence level C

- Nose, axillae, perineum, skin lesions and manipulated sites of the infected patient and all other patients in *"high risk"* units should be screened for carriage of MRSA. Units caring for patients at high risk for suffering serious MRSA infections or with a high proportion of MRSA infections among colonized patients include: intensive care, neonatal intensive care, burns, transplantation, cardiothoracic, orthopaedic, trauma, vascular surgery, renal, regional, national and international referral centres, and other specialist units as determined by the local infection control team and as agreed with the senior clinical staff of the units and relevant hospital management structure.
- MRSA patients can be routinely cohorted with other MRSA patients. Patients with MRSA isolates that are eradicable because of known susceptibility to multiple drugs useful for eradication, should however not be cohorted with those with isolates resistant to these drugs, if eradication will be used as an adjunctive measure.
- Selection of protective equipment must be based on an assessment of the risk of transmission of micro organisms to the patient, and the risk of contamination of the healthcare practitioners' clothing and skin by patients' blood, body fluids, secretions or excretions.
- Gloves must be worn as single-use items. This means that they must be put on immediately before an episode of patient contact or treatment and removed as soon as the activity is completed. Gloves must be changed between caring for different patients (cohort), and between different care or treatment activities for the same patient, moving from a contaminated-body site to a clean-body site during patient care.
- Disposable plastic aprons should be worn when there is a risk that clothing may become exposed to blood, body fluids, secretions or excretions, with the

exception of sweat. They should be worn as single-use items, for one procedure or episode of patient care, and then disposed of as clinical waste.

- Masks should be worn as part of isolation precautions when entering the room of a patient colonized or infected with MRSA, to decrease nasal acquisition by healthcare workers.

4.4.5 Special mouth care (NMDSII item F500)

The following refined PICO matches NMDSII item specifications:

- **Patient characteristics/Population:**
    - Patient receiving chemotherapy or radiation (full body or head-mouth region), special focus on palliative oral care is excluded.
    - Patients are at risk of developing mouth mucosa problems OR Patients who have developed mouth mucosa problems.
    - Patients are adults and receive care in an acute hospital.
- **Intervention characteristics:**
    - Prevention and treatment of mouth mucosa problems: mouth care, rinsing, assessment condition mouth.
    - Brushing teeth or evaluation of teeth as a solitarily intervention is excluded.
- **Comparison intervention:**
- A specific alternative intervention or no intervention
- **Outcome:**
- An evidence based indication of a positive, negative or an absent effect of intervention or a sufficient level of evidence stated to justify an intervention or no intervention

Based on this PICO the following evidence based recommendations can be formulated:

Evidence level A

- The use of Fluoride toothpaste twice a day is highly recommended to obtain an effective prevention and control of dental caries.
- There are different existing oral cleansing agents which can be used as a supplement for oral care. Only few of these agents are recommended based on sufficient evidence. Chlorhexidine mouthwash can be used as a compliment to oral care procedures because it diminishes oral colonisation of micro-organisms.
- Hydrogen Peroxide should not be used on a daily basis.
- Partly absorbed antifungal drugs have a good preventive effect in development of candidiasis and are preferable to absorbed and non absorbed antifungals.
- Benzydamine has a preventive effect in developing mucositis for patients undergoing radio therapy for the treatment of head – neck cancers.
- Benzydamine and dyclonine HCL have a positive effect in treatment of mucositis symptoms for patients undergoing radiotherapy for treatment of head and neck cancers. Hydrolytic enzymes reduce moderate and severe mucositis.

Evidence level B

- The use of the Oral Assessment Guide (OAG) on patients identified as requiring assistance with oral hygiene during routine assessment is recommended.

- Tooth brushing should be the first line of oral cleansing method unless the patient is prone to bleeding, pain or aspiration.
- Foam swabs/brushes with chlorhexidine or toothpaste are only to be used when tooth brushing is contraindicated (e.g. bleeding tendency). Its recommended not to use foam swabs longer than necessary because it is less effective in removing debris and plaque as compared with tooth brushing.
- During and after therapy, a daily assessment of the mouth is needed to discover problems in an early stadium. There is no evidence that any assessment tool is better than other, so the OAG can be used.
- Cryotherapy (providing ice – chips) during the administration period of chemotherapy has a possible positive effect in preventing mucositis.
- Rincing with salt solution preferable 8 to 10 times a day is also advised.
- For patients with 5 fluorouracil chemotherapy with or without radiotherapy allopurinol rinsing 4-6 times per day is preventive for mucositis.
- Other solutions have some effect in prevention of mucositis and can be considered for extra use: calcium phosphate, povidone, zinc sulphate, antibiotic paste or pastilles (moderate effect). Future research is needed to compare these agents with each other as well as to explore the cumulative effect of a combination of agents.
- Oral care protocols have some benefits in reducing severity of mucositis. However, comparison and cumulative effects of agents have not been investigated. Therefore it is not possible to suggest one product over another. Clinical experience, comfort of the patient and cost-effectiveness analyses should be the leading rule in choosing the correct agent.
- Some agents have a limited effect in reducing severity of mucositis. In some cases these products can be considered: calcium phosphate, povidone, zinc sulphate.
- Benzydamine hydrochloride mouthwash, diphenhydramine EMLA, or mixed 1/1 with kaopectate or aluminium hydroxide (Maalox), viscous lidocaine 1:1 with benadryl plus Maalox, steroid mouthwashes are effective in treating a wide range of aphtous ulcers of heightened severity. Side-effects of steroids should be monitored because of inadvertently swallowing.
- The following product combinations are effective in pain relief: topical dyclonine or lignocaine, Diphenhydramine with EMLA or Maalox, viscous lidocaine plus Maalox, chlorhexidine gluconate when aphthous ulcers in a wide range of oral sites aren't accessible using covering pastes.

Evidence level C

- Brushing Teeth with a soft-bristled, small-ended toothbrush at least twice a day, preferable after awakening in the morning and before going to bed is suggested.
- Frequency of brushing is depending on the patient's comfort and status of the oral cavity.
- Appropriately diluted Sodium Bicarbonate, and normal saline mouthwash are useful in certain circumstances like, respectively, dissolving viscous mucous and promotion of granulation and healing.
- Glycerine-based products are contraindicated for use.
- Before radiation- or chemotherapy a dental screening by a dentist or dental team should take place. Treatment of caries and dental disease must be preformed.
- Honey can be considered for extra use in the prevention of mucositis.

4.4.6 Urinary incontinence care (NMDSII item B230)

The following refined PICO matches NMDSII item specifications:

- **Patient characteristics/Population:**
  - Patients suffering from urinary incontinence
  - Adults – Women
  - In a general hospital environment
  - Applicable within the Belgian health care context
- **Intervention characteristics:**
  - Preventive or curative care
  - Including information collection, best choice of materials
- **Comparison intervention:**
  - Alternative intervention, no intervention or procedure
- **Outcome:**
  - An evidence based indication of a positive, negative or absent effect of the intervention
  - Level of evidence to justify the intervention or no intervention

Based on this PICO the following evidence based recommendations can be formulated:

Evidence level A

- Nurses educate and monitor the patient during his stay.
- In case of mild stress incontinence, a follow up on weight reduction can be indicated.
- As a Primary care treatment, diet, behavioral modification, Kegel exercises, environmental modifications and absorbent products are recommended. Most of these interventions can be initiated during a hospital stay.
- Vaginal tampons help to sense the muscles in pelvic floor muscle training (PFME) and prevent incontinence in short-lasting physical strain. A specialized nurse is responsible for supplying the aids and educating the patient.
- A trial of supervised pelvic floor muscle training of at least 3 months' duration should be offered to women with stress or mixed UI as first-line treatment.
- Pelvic floor muscle training programmes should comprise at least eight 8 contractions performed three times per day.
- Pelvic floor muscle exercises should be tailored to be achievable by the individual patient.
- Ensure women are performing PFME exercises correctly.
- Pay particular attention to women with antenatal and postnatal urinary incontinence in providing advice and PFME instruction, incorporating at least two individual instruction sessions into the program.
- PFME programs should be multi-faceted with a number of components, rather than supplying printed information only.
- Bladder training is strongly recommended for management of urge UI.
- A period of caffeine reduction is recommended alongside bladder training.

- Prompted and timed voiding toileting programmes are recommended as strategies for reducing leakage episodes in cognitively impaired women. Initiate an individualized prompted voiding schedule based on the client's toileting needs, and as determined by a 3-day voiding record. The identification of an individual voiding pattern can promote the highest level of continence for the individual while reducing the time required to toilet, including the time of the care provider. There is strong evidence that prompted voiding reduces the frequency of incontinence in individuals who can initiate voiding when prompted. Literature also suggests that although prompted voiding does not require expensive equipment to implement, the consistent availability of a care provider or staff to provide the prompted cues is a factor in the success of prompted voiding. Several studies conclude that prompted voiding is easy to learn, but requires personal dedication and consistent application of the protocol.
- Urinary incontinence in women can be effectively managed ire general practice with fairly simple treatment.

Evidence level B

- Detecting the symptoms, severity and type of incontinence can help to decide the best cost-effective solution for the patient.
- Pelvic floor training is based on an awareness of the intervals between voiding and voiding frequency (including a voiding diary). It is often combined with pelvic floor exercises. The technique is often used in women aged over 55. It improves stress urinary incontinence and urge incontinence.

Evidence level C

- Use paper diapers: they have 2-3 times the fluid capacity of cloth diapers.
- Bandages, diapers, urinals, and plastic bed sheets prevent leaking.
- All women with urinary incontinence should be initially offered nonsurgical therapy since it large percentage will obtain satisfactory results.

4.4.7 Urinary catheter care (NMDSII item B250, B300)

The following refined PICO matches NMDSII item specifications:

- **Patient characteristics/Population:**
  - Patients suffering from urinary incontinence
  - Adults – Women
  - In a general hospital environment
  - Applicable within the Belgian health care context
- **Intervention characteristics:**
  - Incontinence Pads or Absorbent Pads or Diapers
  - Including information collection, best choice of materials
- **Comparison intervention:**
  - Timed voiding
  - Pelvic Floor Exercises
- **Outcome:**
  - An evidence based indication of a positive, negative or absent effect of the intervention

Based on this PICO the following evidence based recommendations can be formulated:

Evidence level A

- In certain cases other catheterization methods can be useful as an alternative of an indwelling bladder catheter: urine collection using a condom catheter, sub pubic catheterization, and intermittent catheterization.
- To assure a free urinary flow: the risk of torsion of the catheter and collector tube should be avoided. The collector is emptied on a regular basis using a separate device for every patient. Contact between sterile and non sterile parts has to be avoided. A dysfunctioning or obstructed catheter has to be irrigated and replaced if necessary. The collector is continuously placed below bladder level.
- A catheter has to be inserted using aseptic techniques, using sterile materials. A permanent catheter is fixed in a comfortable way aimed at avoiding catheter movement and/or urether traction.
- The drainage system is kept sterile and closed permanently. Catheter and collector aren't deconnected, unless irrigation is indicated.
- In case of a leak, deconnection or error in aseptic handling, the collector is replaced in an aseptic way after desinfection of the catheter – collector connection.
- The meatus should be washed daily with soap and water.
- Bladder instillations or washouts must not be used to prevent catheter-associated infection.
- Hands that are visibly soiled, or potentially grossly contaminated with dirt or organic material, must be washed with liquid soap and water. Hands must be decontaminated, preferably with an alcohol based handrub unless hands are visibly soiled, between caring for different patients and between different care activities for the same patient. All healthcare personnel must wear a new pair of non-sterile gloves before manipulation of the system. The hands are washed immediately before and after each manipulation of catheter entry point or drainage system.
- Sterile high volume syringes and a sterile irrigation solution are used an disposed of. Irrigation is performed in a aseptic way.
- Only caregivers (healthcare workers, family members or pateints themselves) with knowledge and skills in aseptic techniques in placing and caring for a drainage system may manipulate its components.
- Carers and patients managing their own catheters must wash their hands before and after manipulation of the catheter.

Evidence level B

- If there is a predictable elevated risk of obstruction (e.g. caused by post operative bleeding of prostate or bladder), a reflux of irrigation needs to be avoided. A closed system of continuous irrigation can be used to avoid this type of obstruction. To clear an obstruction caused by urinary cristalloids, mucus or other causes an intermittent irrigation method can be used. Continual bladder irrigation with antimicrobial agents hasn't proven effective. A systematic application of this technique to prevent urinary infections should be avoided. If the urinary flow can't be maintained unless frequent irrigations are applied, the catheter has to be replaced immediately if it seems likely that the catheter itself contributes to the obstruction.
- Choice of catheter among the prescribed method: The number of the catheter gives its circumference in millimetres. The diameter of the catheter is roughly the circumference divided by 3. Silicone and PVC are the most suitable materials for long-term catheterization as they cause the least tissue irritation. A silicone

catheter (size 12–14) is preferable. A PVC catheter (with a larger internal diameter) is most practical if the urine is bloody and flushing of the bladder is necessary. If the urine is bloody a Ch 16, PVC catheter can be used. Many catheter-associated problems can be avoided by selecting a closed catheter system with a small size catheter (14 to 18 French with a 5-cc balloon), following manufacturer's recommendations for inflation/deflation, maintaining a closed system, securing the catheter, and properly positioning the drainage bag.

- Installing the urinary catheter or indwelling devices and follow the drainage: Material aids include gloves, a sterile underground, indicated antiseptic solution for a peri uretral washing and a unique dose of lubrifiant creme. The catheter has to be of the smallest measure as possible while garding an adequate urinary flow. This choice minimizes uretral traumata.
- A permanent catheter doesn't have to be replaced at fixed arbitrary intervals.
- A two daily iodium based antiseptic cleansing in combination with daily hygienic care hasn't proven beneficiary.
- Hands must be decontaminated immediately before each and every episode of direct patient contact or care and after any activity or contact that could potentially result in hands becoming contaminated.
- Continuous education of caregivers is indicated. Adequate techniques and detection of complications should be highlighted.

Evidence level C

- Identifying needs of patients who no longer need indwelling catheters and discussing appropriate catheter alternatives: In collaboration with the physician and because of the complications of long-term indwelling catheter usage, periodic assessment and voiding trials should be used to determine the continued need for a catheter.
- Based on the prescribed protocol and prior to inserting the catheter or prior to replacing the catheter, instill 2 percent lidocaine jelly (if immediately available) into the urethra and wait 2 minutes, if possible. Catheterization can exacerbate autonomic dysreflexia. The use of lidocaine jelly may decrease the sensory input and relax the sphincter to facilitate catheterization. The peak effect of lidocaine jelly is between 2-5 minutes. Exercise clinical judgment regarding elevated blood pressure; immediate catheterization may be necessary.
- If the catheter appears to be blocked, gently irrigate the bladder with a small amount (10-15 cc) of fluid, such as normal saline at body temperature. Irrigation should be limited to 5-10 ml for children under 2 years of age and to 10-15 ml in older children and adolescents. Avoid manually compressing or tapping on the bladder.
- If the catheter is not draining and the blood pressure remains elevated, remove and replace the catheter.
- If difficulties arise in replacing the catheter, consider attempting to pass a coude catheter or consult a urologist. A coude catheter may be useful if there is an associated bladder neck obstruction.
- Monitor the individual's blood pressure during bladder drainage.
- Sudden decompression of a large volume of urine would be expected to normalize blood pressure. However, this may cause hypotension if the individual has already been given pharmacological agents to decrease the blood pressure.
- The urinary flow shouldn't be interrupted unless for medical reasons.
- If the individual has an indwelling urinary catheter, check the system along its entire length for kinks, folds, constrictions, or obstructions and for correct placement. If a problem is found, correct it immediately.

- Always insert a catheter aseptically, use a closed drainage system, and properly maintain the catheter.
- Inject 20 ml gel into the urethra of men (and somewhat less for women). Use preferably a gel containing a local anaesthetic.
- Both gel injection and insertion of the catheter should be performed gently and slowly.
- Fill the balloon only after making sure that both the tip of the catheter and the balloon are in the bladder: the urine flows freely, or if the bladder is empty, saline solution injected into the catheter flows in easily.
- In long-term catheterization the balloon should be filled with 5% saline or glycerol solution.
- The catheter must not be pulled downwards by gravity (use a thigh bag).
- Practices such as routine catheter irrigation should be avoided.
- Removal of catheter blockage is preventive for renal disease.
- Catheter Maintenance: Leave the closed system alone! Wash the urethral orifice with an antiseptic solution (e.g. 0.01% chlorhexidine). Appropriate maintenance minimizes infections. Meatal cleansing with antiseptic solutions is unnecessary.
- To reduce the incidence of CR-UTI by maintaining a closed system, all LTC must be connected to a sterile closed drainage system or valve.
- To reduce the incidence of CR-UTI caused by blocking, all newly catheterised patients should have a patient record that documents the integrity of the catheter at first change and adjustments made to their change schedule accordingly.
- Make sure you use the correct technique when using indwelling devices as it is vital to reduce the risk of patients acquiring infection. 80 per cent of urinary infections can be traced back to indwelling urinary catheters. These infections arise because catheters traumatise the urethra as well as providing a pathway for bacteria and other organisms to enter the bladder. The longer such catheters are in place, the higher the risk of infection.
- For drainable catheter systems used by primary care patients, daily bag cleaning with a diluted bleach solution (1:10) is effective in reducing bacterial counts to negligible numbers.
- Urinary white blood cells are the best indicator of urinary tract infection.
- All patients should have a patient record that documents the reason for catheterisation, type of catheter, catheter insertion, changes and care.
- Ensure that all healthcare personnel are trained and competent in urinary catheterisation, Evidence of healthcare personnel training and annual assessment of competence.
- Educate patients and manage pain: Current practice in pain management and care of a patient with an indwelling urinary catheter is evaluated against best practice suggested by recent research evidence. A multidimensional and multidisciplinary team approach is followed to alleviate pain and promote independence in catheter care. Imagination and a positive attitude led to greater patient comfort and dignity. To ensure patients and carers are informed and educated about catheter management; all patients and carers are aware of the need to decontaminate their hands, keep the system closed and seek professional help when they suspect clinical infection.

4.5 DISCUSSION AND CONCLUSIONS

The selection of interventions and the literature review methodology could be performed adequately within an evidence based framework. Not all selected interventions within a defined PICO context are equally applicable. An Artificial Ventilation PICO for example was also tried. It didn't generate many evidence based indications, due to a lack of scientific research for the chosen type of ventilation and the existence of inconsistent evidence based findings. Other PICO's led to sufficient information. Generally, it was surprising that so many guidelines and systematic reviews are already developed with a main focus on evidence based nursing. Currently, not many of these sources are used in hospital nursing practice. The first potential issue, unavailability of E.B. knowledge in nursing, isn't confirmed. A set of indications and contra indications can be extracted from the evidence based literature recommendations. E.B. guidance about the execution of the selected interventions is also available. One limitation to this conclusion is the use of 'availability of an E.B. body of knowledge' as a criterion for selecting the investigated nursing interventions. These interventions were already at the start considered evidence rich. This may be untrue for other nursing interventions as described in NMDSII.

A second observation concerns the level of evidence. There is already a considerate amount of level A and level B evidence available in nursing research. However, much of the EBN guiding recommendations are still of a level C. Most of these are built on strong expert consensus and are logically indispensable in providing high quality nursing care. For example, there is no A or B level evidence available concerning regular systematic skin inspection in the prevention of pressure ulcers. This clinical assessment is of such an importance as a first cornerstone to build on in further E.B. treatment, that it can't be dismissed. There is a rationale to include C level knowledge if it's sufficiently theoretically grounded and evaluated by an E.B. review expert panel that reaches consensus.

5 FEASIBILITY OF AN EBN – RULESET ON A DATABASE LEVEL

5.1 INTRODUCTION

Since the 'to err is human' report of the IOM in 2000 on the disturbingly high incidence of adverse events in general healthcare, quality of healthcare is becoming an explicit priority. Overuse, under use and misuse of healthcare interventions lead to avoidable morbidity and mortality [99, 100, 101]. The current healthcare finance system isn't quality improving. A focus on quantity, length of stay, productivity and cost control is paramount. To the contrary, avoidable complications and a low quality healthcare in general lead to a higher hospital reimbursement [102, 103]. Booming during the previous years, in the period 2003-2007, a movement aimed at relating healthcare finance to quality of care is gaining ground internationally. 'Pay for performance' or 'pay for quality' applications provide for quality inducing incentives and the removal of disincentives by financial means. Already hundreds of experimental projects have been implemented in countries such as the US [104], the UK [105], Canada, Australia, New-Zealand, Haiti, Nicaragua and Spain. The first effect studies indicate a mixed to positive result [106, 107, 108]. However, there is a great lack of evidence concerning the appropriate design of such a system. Moreover, incentives are mostly aimed at quality care as influenced by physician behavior. Interdisciplinary care, including nursing interventions, is rarely considered. Therefore a first evaluation of potential of such a system on a nursing care level is indicated. As discussed in the previous chapter, EBN guidance for specific nursing interventions is available. It is however unclear to which degree practice can be evaluated by integrating an EB – rule set of appropriate care into current Belgian clinical information systems.

5.2 RESEARCH QUESTION

This chapter is aimed at the following research question:

Is it possible to evaluate the appropriateness of nursing care based on using an EBN rule set?

5.3 METHODOLOGY

Pressure ulcer prevention has been selected as nursing intervention to test EBN – rule set applicability.

The first step towards an evidence based rule set is the translation of the evidence based pressure ulcer recommendations (section 2.4.) to a clear stepwise framework capturing the decisional logic behind clinical practice. In evidence based methods this is often done by means of the construction of decision trees and algorithms. Most evidence is only available in a text-based format. The migration to IT databases requires translating text into intermediary representations. One type of representation that is readily adaptable for computerization is a linear algorithm [109]. The process includes: defining applicability criteria, identifying entry points, defining decision points, defining actions and creating a linear algorithm that links decision points and actions. This process is also known as 'knowledge specification' [110]. The main hypothesis is that the core recommendations of guidelines can be represented as a decision tree, independent from the encoding model (e.g. SAS input). This tree is composed with steps (root, nodes and leaves).

Figure 9 presents the main algorithm, deducted from evidence based recommendations. The level of evidence is also presented for each component in the decision tree. The results, as discussed in the next section, are also projected onto the algorithms.

A patient moves through this main algorithm in the following way: After admission the risk for developing pressure ulcers should be assessed within two hours after admission. This risk assessment is presented in a sub algorithm (Figure 10). Use of a systematic assessment

tool such as Braden or Norton, systematic skin inspection and self inspection by the patient if possible, are performed to assess primary and secondary risk factors. If there are no risk factors present, then interventions are not indicated. This is represented by the first exit point in the main algorithm. If only secondary risk factors are present, no intervention is indicated, but regular risk reassessment is needed. This is presented by the upward loop in the main algorithm. If primary risk factors are detected, the patient is considered to be at risk. If primary and secondary factors level A are present in combination, the patient is considered to be at high risk. The only difference in interventions between risk and high risk concerns the preferred use of a dynamic system mattress over a static system. Afterwards both groups merge in the decision tree.

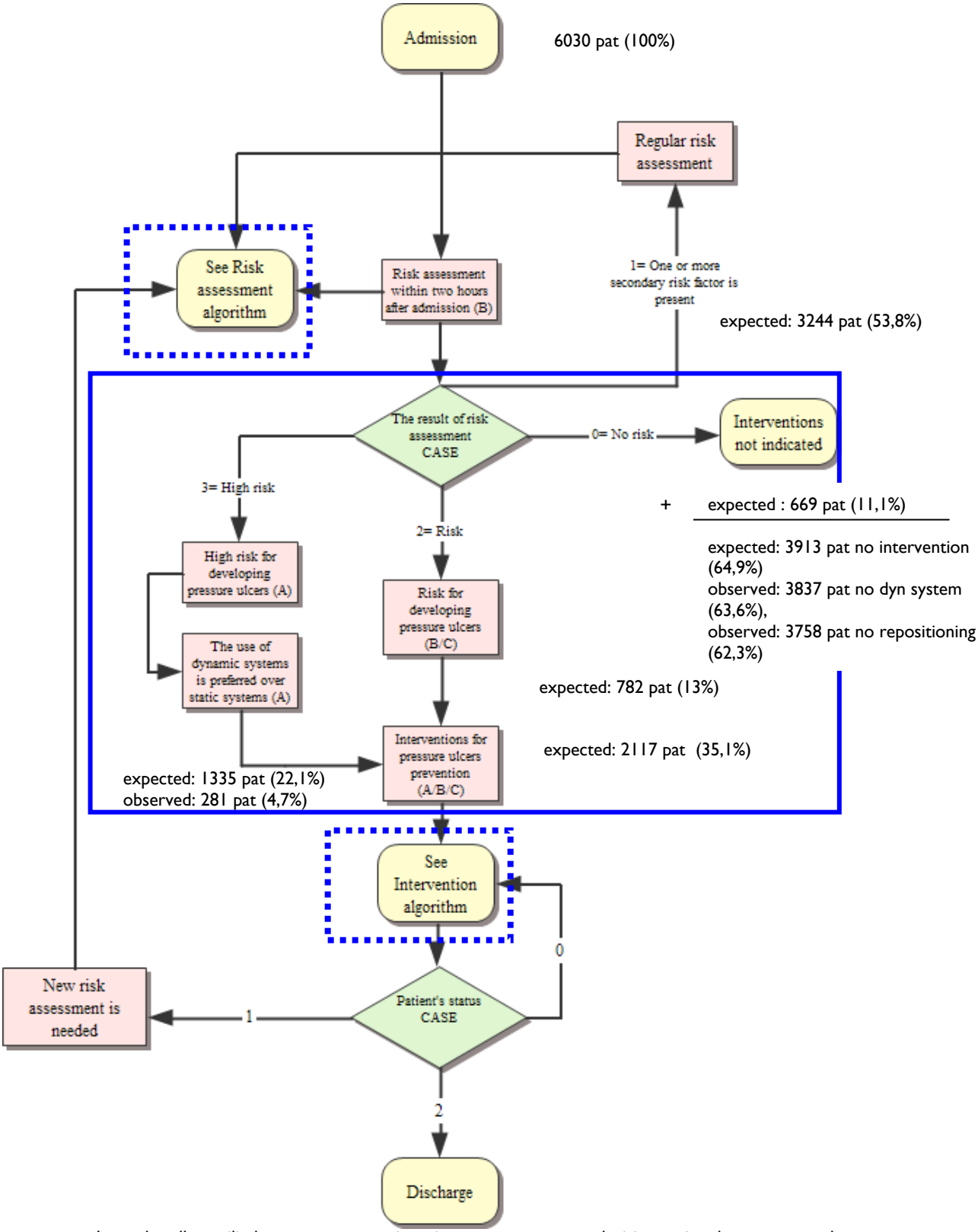
The intervention sub algorithm (see Figure 11) explains the further interventions. A distinction is made between pressure ulcer specific interventions and basic care intervention adaptations.

The basic care adaptations prescribe points of attention in mobility, activity and hygienic care for all patients who are at risk. Additional modifications are needed if the patient is wounded or incontinent, if care aids are present which may stimulate pressure ulcer development and if the patient has an insufficient nutritional and/or fluids balance.

Pressure ulcer specific interventions consist of three main components: the use of a special mattress, changing patient positions and patient education. If repositioning is impossible due to clinical contra indications, a dynamic system mattress is needed. If repositioning is possible, the frequency depends on the presence of a special mattress or not. If a special mattress is present, the patient should be repositioned every four hours. If no special mattress is present, the patient should be repositioned every two hours.

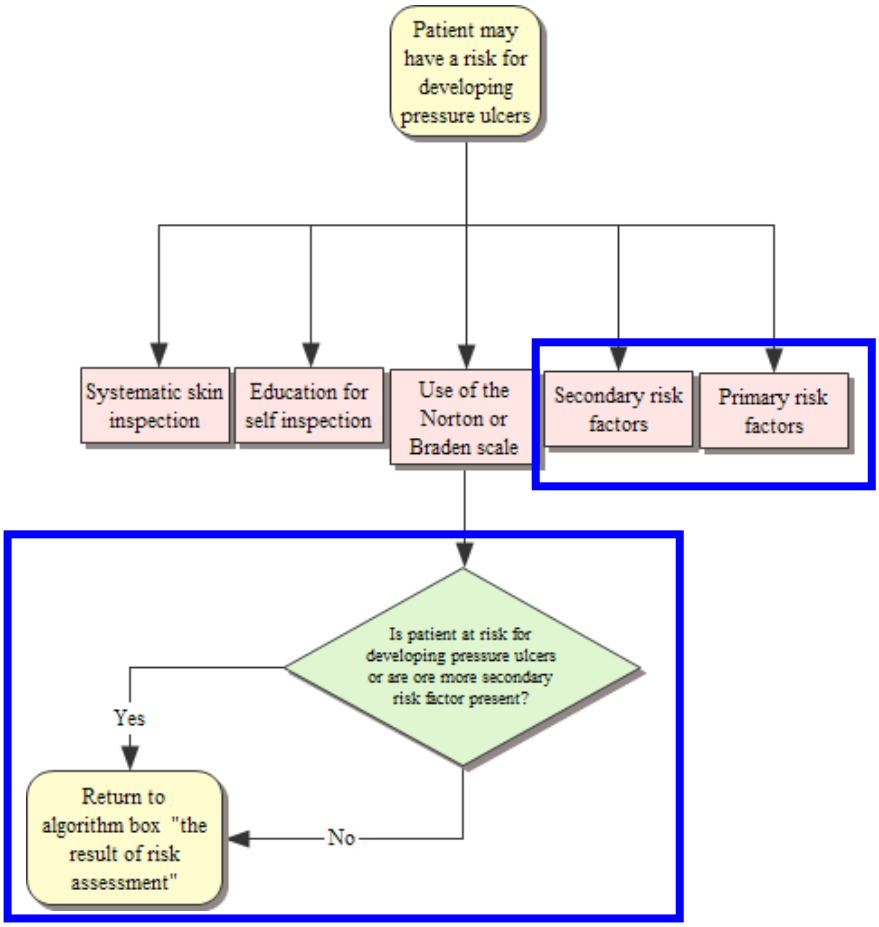
Back in the main algorithm at the bottom one sees that a new risk assessment is needed if the patient status changes concerning one of the primary or secondary risk factors. This is presented by the high upward loop. If the condition of the patient remains unchanged, reassessment should take place every 72 hours. This is presented by the small upward loop. The algorithm forms a continuous cycle of diagnosis and intervention.

Figure 9: Main pressure ulcer prevention algorithm



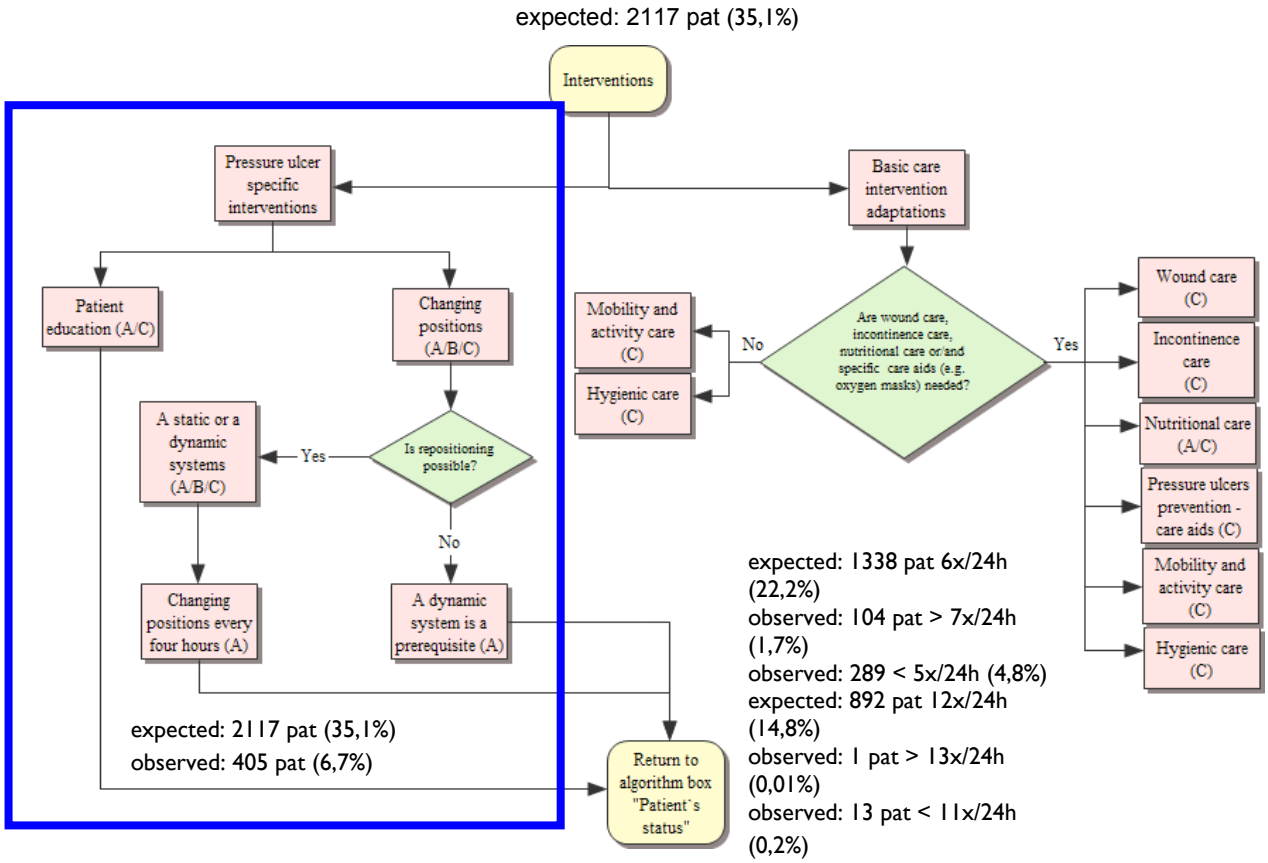
Legend: yellow cilinder = entry or exit point, green square = decision point, brown rectangle = intervention, full blue border = available within national minimal databases, dotted blue border = partially available within national minimal databases, expected and observed are always calculated based on the total number of 6030 patients. A, B and C represent levels of evidence.

Figure 10: Sub algorithm for risk assessment



Legend: yellow cilinder = entry or exit point, green square = decision point, brown rectangle = intervention, full blue border = available within national minimal databases, dotted blue border = partially available within national minimal databases, expected and observed are always calculated based on the total number of 6030 patients. A, B and C represent levels of evidence.

Figure 11: Sub algorithm for preventive interventions



Legend: yellow cilinder = entry or exit point, green square = decision point, brown rectangle = intervention, full blue border = available within national minimal databases, dotted blue border = partially available within national minimal databases, expected and observed are always calculated based on the total number of 6030 patients. A, B and C represent levels of evidence.

The next step in database rule set development concerns the availability check of all of these decision tree components in the national databases of minimal registered data, i.e. HDDS and NMDSII. The parts that are retrieved in the databases are indicated in the algorithms by a blue rectangle (dotted if only partially available).

The assessment part of the algorithm can partially be linked to minimal data. The execution of assessment with a Braden or Norton tool and skin inspection is out of database scope. Thus timeliness of this assessment can also not be checked. However, the primary and secondary risk factors can be retrieved to a large degree based on ICD 9 codes in HDDS and certain NMDSII items. So risk can be assessed to a certain degree on database level. This means that the assessment itself as intervention can't be evaluated, but the appropriateness of subsequent interventions can be evaluated based on risk outcomes. Relevant ICD 9 and NMDSII codes are clustered towards primary or secondary risk factors.

Primary risk factors are captured by codes concerning patient immobility such as patient installation, consciousness, amputation, paralysis, use of complex devices, etc. Secondary risk factors are based on codes about nutrition, incontinence, sensory abilities, diabetes, neuropathy, ulcer, erythema, pain, fracture, obesity, stroke, dementia, COPD, myocardial infarction, anemia, intensive cancer treatment, wound presence and high age.

The intervention algorithm can also partially be evaluated on a database level. Basic care adaptations are very subtle and detailed. They are beyond the database scope. However, pressure ulcer specific interventions can be evaluated. Mattress systems, changing positions and patient education are largely present in NMDSII, with some limitations discussed below.

The last part of the algorithm, about reassessment and its timeliness, falls outside the database scope.

A HDDS – NMDSII merged dataset was created, including original ICD 9 codes. NMDS-II information was collected during the pilot study of actualisation NMDS-II and consists of 117.395 inpatient days from 66 Belgian hospitals. Each hospital participated in the test with a minimum of 2 and a maximum of 5 nursing wards. A balanced sample was obtained for the following medical specialties: geriatrics (index G), pediatrics (index E), intensive care (index I), chronic illness (index SP), maternal services (M), and general internal medicine (index D, H*) and general surgical procedures (index C, H*)

From these hospitals the MKG data for the corresponding hospitalization units and registration period (year 2003 semester 2 and year 2004 semester 1) was collected. From 59 hospitals the quality of the data was sufficient to create a dataset of DRG's using the apr-drg grouper from 3M.

20990 records on a hospital day of stay level could be coupled assuring optimal data quality. Afterwards data were aggregated on a individual patient level concerning 6030 patients. The decision tree was programmed using SAS version 9.1. A regular type of IT programming in logical if – then statements was used. In this way, the rules within the decision tree on paper become applicable on large databases and other kinds of spreadsheets. The rules can afterwards also be rechecked efficiently by running the same program again after a period of time or on other similar data. The program was fine tuned systematically towards its optimal use. Updating the system can be done fairly easily by modifying a statement.

5.4 RESULTS

2117 of the 6030 patients (35, 1%) have a primary risk factor on one or more days during their registered stay; 2511 (41, 6%) patients have a secondary risk factor level A; 4786 (79, 4%) patients have a secondary risk factor level B or C. Combinations of different types of risk factors within one patient stay are possible, even simultaneously.

These factors translate to a risk classification in a following way:

- 669 of 6030 patients (11, 1%) have no risk during their stay. There is no primary, nor secondary risk factor present. Interventions are not indicated for these patients.
- 3244 (53, 8%) patients have no primary risk during their stay, but one or more secondary factors present. For these patients only regular reassessment is indicated.
- 782 (13%) patients have a primary risk factor on one or more days during their stay. Preventive interventions are warranted.
- 1335 (22, 1%) patients have a primary risk factor in combination with an A level secondary risk factor on one or more days during their stay. Next to standard preventive interventions the use of a dynamic system special mattress is indicated.

So for a total of 2117 of 6030 patients (35, 1% of total) the use of preventive interventions is indicated.

All of these 2117 patients should receive patient education about pressure ulcer prevention. 1335 of them should be placed on a dynamic system special mattress as already mentioned. These same patients plus the ones currently placed on another special type of mattress equal 1338 patients (22, 2% of total). They should be repositioned six times a day. The other 892 patients (14, 8% of total) should be repositioned 12 times a day.

It is possible to compare the justified figures above with the care as actually registered in minimal data. First, we can assess the level of over care in patients who didn't needed preventive interventions based on registered data, mostly NMDSII. Of these 3913 patients (64, 9% of total) with risk = 0 or 1 (only reassessment), 76 (1, 3% of total) did lay on a special mattress, although this wasn't indicated. And 155 (2, 6% of total) did receive repositioning, although this wasn't indicated.

Secondly the level of under care is assessed. Of the 1335 patients (22,1% of total) who should be placed on a dynamic mattress system, 1054 patients (17,5% of total) didn't receive this warranted care. Only 281 of them (4, 7% of total) did receive this appropriate care. Of the 2117 patients (35, 1% of total) who should receive patient education 1712 patients (28, 4% of total) didn't receive this warranted care. Only 405 of them (6, 7% of total) did receive this appropriate care.

In assessing the adequacy of repositioning a buffer was used as an interval [-1; +1]. For example if a patient should be turned six times, five or seven were also considered correct. The same holds for twelve times turning.

Of the 1338 patients (22, 2% of total) who should be turned six times a day, 104 patients (1, 7% of total) where turned more than seven times a day. This also constitutes a type of over care. The average turning range during their stay goes up to 14 times for some of these patients.

In contrast, 289 of the 1338 patients (4, 8% of total) were turned less than five times a day. For several dozens of these patients this means only one or two times a day on average during their stay.

The same analysis can be performed for 12 times turning: Of the 892 patients (14,8% of total) who should be turned twelve times a day, only one patient (0,01% of total) was turned more frequently than 13 times a day as a type of over care. And only 13 of the 892 patients (0, 2% of total) were turned less than 11 times a day.

5.5 DISCUSSION AND CONCLUSIONS

The pressure ulcer prevention example shows that it is feasible to construct an evidence based rule set implemented on a database level. Not all, but already some important aspects of care can be compared in this way concerning reality versus appropriateness. Although one has to take into account the high amount of time and resources that have to be reserved to construct a decision tree, review database variables (ICD 9), recode original database variables, and to program the decision tree.

Unavailability of data limits the potential applicability of an evidence rule set on a national database level. But the parts that are present can be programmed in logical if – then statements, leading to interesting results.

In interpreting the results above one always has to keep in mind that it concerns minimal data and not direct care observations. These minimal data hold a certain degree of bias. Underscoring and over scoring are both possible in registration. The observed under use can be influenced by under coding. Coding behavior is also partially driven by the goal and application of the data registration afterwards. The data as originally collected are used in the current Belgian financing system. So one would expect a degree of up coding. The results concerning under care are even more striking, taking this potential bias into account.

In general there appears to be a great degree of under care present in the prevention of pressure ulcers. The figures about the use of a dynamic system special mattress and patient education are startling. In repositioning of patients on a special mattress there is also a considerate degree of under care present. Over care in preventing pressure ulcers is not common, based on a minimal data level analysis.

It is also apparent that at every level of the algorithm the level of evidence is a mix of A, B and C degree. Without certain B and/or C level components the algorithm falls apart. So, an exclusive use of level A evidence for these nursing interventions would lead to under care in clinical practice. This gap can be closed only in long term by generating more A level evidence. A level A, (A and B) or (A, B and C) rule set could theoretically be deducted. A strict level A rule set is expected to miss many clinically relevant elements, because of the lack of strong evidence generation in the field of nursing. A broad (A, B and C) rule set takes also common practice into account, but lacks sufficient evidence to justify all elements. An (A and B) rule set can present the middle ground.

The future feasibility of rule set construction will depend mostly on the availability of data in the NMDS-II, HDDS and other data sets for further investigation. Also the coding characteristics (e.g. nominal, ordinal or continuous) influence the applicability of specific elements within the model. Certain NMDS-II items consist of more than one nursing intervention. This makes one on one relationships more difficult to discern. The data sets can also be used for checking of epidemiologic research, as presented in the form of the indications and contra indications. In a more general manner registered data can be used to investigate all sorts of relationships predicted in evidence based literature. Limitations, such as the coding quality, have to be taken into account.

In the present study only nine NMDSII interventions were specifically reviewed on evidence based practice, and a rule set is constructed concerning one nursing intervention. In future research more relationships can be investigated by combining rule sets. The pressure ulcer prevention recommendations for example, also have implications for hygiene care, incontinence care, wound care, etc. This interdependence is found in most of the selected interventions.

6 NURSE STAFFING NEEDS IN A GENERAL HOSPITAL ON A PATIENT CASE LEVEL

6.1 INTRODUCTION

The actual financing system for nursing activity in Belgium consists of a basic and a supplementary part in budget allocation to nursing wards. An appropriate volume of care is attributed to hospitals per APR-DRG in terms of patient days. The term 'appropriate' can be misleading. It is restricted to a quantitative evaluation of how many days should be expected, based on the national average length of stay per APR-DRG. The revenue in 'price' terms as a function of this expected care volume is determined in the following way. The basic part of nurse staffing financing in budget allocation per nursing ward is based on a fixed minimal quantitative staffing norm per medical specialty. Such norms determine the minimal number of FTE's that have to be available during 24 hours of care for each specialty. The supplementary part of nurse staffing financing is allocated over hospitals based on a 1 to 10 deciles ranking of hospitals. Criteria differ with medical specialty. However, for general hospital care the following main criteria drive the ranking system: Firstly, the relative reimbursement value of performed medical interventions as a total of fee for service bills; Secondly, the value of the mean NMDS – weights per patient day as measured by the national Nursing Minimal Data Set. Before 2007 this set consisted of the registration of the presence or frequency of 23 types of nursing interventions. The set was registered four times a year during 15 days on each Belgian nursing ward.

The mean NMDS – weights, as one of the ranking criteria, are calculated in a complex manner. Multidimensional scaling projects every nursing ward on a national 'map' within a dependent – independent care dimension and a basic – intensive care dimension. Every nursing ward is positioned within one of 28 care zones on the map taking into account this nursing profile differentiation. The process determines 28 clusters of nursing wards. The cluster in which a nursing ward falls has a unique NMDS weight. This weight is an indicator of the zone specific staffing characteristics, as a combination of a staff qualification index and a staff quantification index (FTE/patient day).

The need for change of this system is clear. In 2006 NMDS was thoroughly updated towards a system of 79 nursing intervention items. NMDSII is the result of broad qualitative sector participation and a statistical quantitative reconfiguration of the system. It is up to date with current nursing practice. It is based on NIC as an international nursing intervention 'language'. And it is a much more accurate representation of what nursing care incorporates in all its different dimensions when compared to the previous version of NMDS. Since NMDSII is implemented nationwide in 2007, it was also necessary to adapt the supplementary financing part of nursing care in general hospitals as it is partly based on NMDS data. This necessary change can be considered as an opportunity to correct other shortcomings in the current supplementary financing system. The current system lacks in one very important aspect: the NMDS – weighting as financial driver is based on a historically determined staffing qualification and quantification per care zone. There is no transparent relationship with nursing care needs which result from patient care needs. The study part described in this chapter is aimed at redirecting the supplementary part financing system from actual towards justified staffing needs as a key criterion for resource allocation. One of the main challenges in this endeavour is capturing the whole patient care context and its inherent complexity of nursing interventions and interrelations between interventions within the determination of justified staffing needs. In these bottom-up approaches are more sensitive to patient demands [111]. They should drive further top-down planning and financing of nurse staffing time allocation.

Fagerström & Rainio (1999)[112] described the optimal level of nursing care intensity as a balance between the patient's needs for care and the number of nurses available to provide the care. Several authors stress the importance of linking allocation of nursing time to a qualitative system, based on the results of nursing care. The result is the important thing, not simply the time required. It is not enough that patients get the right amount of care

quantitatively. Nursing staff are also responsible for providing patients with nursing care of good quality.

6.2 RESEARCH QUESTION

This study part is aimed at the following research questions:

Is it possible to make a transfer from actual towards justified nurse staffing needs taking into account nursing care requirements to ensure quality of care?

Which are the estimated nursing time needs with regard to specific patient cases to ensure quality of care?

6.3 METHODOLOGY

Several methods can be used to assess nursing time needs: time-and-motion studies, work sampling, and subjective evaluation are the most commonly used [113]. Traditional time studies have been criticized in recent years [112]. It is difficult to fully comprehend nursing care by means of time studies as, by its nature, it is complex and multidimensional. Traditional time studies are also expensive and time consuming. Results of time studies in one ward are not thought to hold true for another ward. The need for staff is also not of a linear nature. Nurses often 'multitask' by doing more than one activity at a time [114]. Both time-and-motion and work-sampling methods are subject to the Hawthorne effect, which consists of workers changing their habitual work pattern when observed [113].

Self reporting is a good, low-cost means of quantifying time allocation by nursing care staff [115]. Nursing care itself starts from personal estimates. Therefore the methods of assessing the need for staff are also personal by their nature [112]. Previous methods add a scientific and statistical dressing to what is essentially professional judgement [116]. At some point all these methods rely on a nursing judgement. This is the result of information processing within the limitations of available knowledge. Valuing the professional judgement of nursing professionals validates their unique and highly desirable contribution to the health care economy [114].

Subjective evaluation usually takes the form of interviews or questionnaires [113]. These are subject to personal biases such as participants' problems with memory, selective recall and correct question interpretation [114]. This studies method is therefore based on a Delphi study in which a panel of charge nurses rated nurse staffing demands for specific patient cases. It is an important method for achieving consensus on issues where none previously existed [117]. Problems with memory and selective recall are prevented. A Delphi approach is a validated technique with known advantages and disadvantages. It is a method to obtain the most reliable consensus of opinion. The non random selection of 'experts' has to be taken into account. Anonymity and a sufficient number of rounds enhance the process [118]. Recommendations as presented by Hasson et al (2000)[119] were followed. The process is explained step by step in subsequent subsections. See also Greatorex & Dexter (2000)[120] for further in depth discussion of the Delphi method.

6.3.1 Creation of patient cases

6.3.1.1 Case selection and participation

At first real patient cases had to be written. Later on these cases are used to assess variability in nursing care needs and to investigate the relationship with staffing needs and thus also financial needs. One hundred cases would be sufficient to map variability in nursing care needs: a minimum of 40 of medico – surgical wards, 20 of paediatric wards, 20 of geriatric wards and 20 of intensive care units was aimed for. The three latter groups are isolated, because each of them entails a very unique case mix, distinct from a regular nursing unit. Other very specific clinical settings, such as a dialysis unit, chronic disease and rehabilitation (SP index) aren't part of the current research project. Specialties such as oncology, cardiology, general surgery, internal medicine, urology, nephrology, pneumology,

gastro enterology and haematology are treated within one group: the medico – surgical wards. Further distinction would be unfeasible, considering the number of cases that are required for each subgroup.

The case construction is based on real patient cases, as encountered in Belgian general hospitals. Patient records in combination with additional information from nurses, involved in the specific care delivered, are the basis for case construction. A case describes the whole of nursing care delivered for a specific patient, during one day of stay (24 hours). NMDSI and NMDSII are also coded for each selected patient case. All information is obtained by way of nursing unit visits. Direct record reviewing and continuous contact with involved caregivers ensure the validity of the constructed cases. After case construction, an additional caregiver feedback warrants a genuine description of care as it was rendered in practice.

To obtain sufficient data all 66 hospitals, that participated in the NMDSII test phase in the framework of the NMDS Actualisation Research Project (2002 – 2006) (listing in appendix 4), were invited to provide patient records. Following inclusion criterion was applied: the charge nurse has a 5 year experience as a nurse and a 1 year experience as charge nurse. 16 Dutch speaking and 22 French speaking hospitals offered to participate.

The choice to include only hospitals that already participated in the NMDS-II test phase is supported by the following arguments:

- Central points of contact were already in place. The NMDS coordinators were used to managing the internal research aspects.
- A thrust worthy relationship of mutual collaboration had already been established during the previous years. This culture could be continued.
- It concerns a sufficient number of hospitals, without adding additional newcomers.
- The representative ness for all Belgian general hospitals in terms of number of beds, teaching status (univ./non univ.) and ownership (public/private) was already confirmed in the previous research project. In geographical terms all Belgian regions and provinces are covered.
- An additional demand in surplus of patient record provision concerned the scoring of NMDSI and NMDSII for each described patient day of stay. In these hospitals was sufficient experience with NMDSII to ensure proper coding.
- For practical reasons, one exception is made for the production of ICU cases: O.L.V. hospital in Aalst didn't participate in the test phase of NMDS-II.

Most hospitals suggested more than one nursing unit to participate. The 16 Dutch speaking hospitals offered input of 44 nursing units and the 22 French speaking hospitals gathered cases in 46 nursing units. Since the number of cases to construct was limited to around 100, a further reduction and selection of hospitals and nursing units was necessary.

The data from the patient records were gathered in several different ways:

- The CHU research partner had the opportunity to collect data by students as part of their internship for obtaining their 'Master in Nursing' title. As ten students participated in data collection, it made a broad selection of multiple units possible.
- For CZV research partners three Master students collected part of the necessary data within the framework of their thesis project. One student worked at an ICU ward in O.L.V. hospital Aalst. This is why O.L.V. Hospital Aalst was included despite not have been participated in the NMDS-II test phase. In addition two job students, who also participated in the Master studies, were hired for data collection. The UZ and CZV research members also obtained part of the data themselves. To ensure feasibility a lesser number of Dutch speaking hospitals was sought. In this way more patient record reviews could be combined within one hospital visit.
- In four Dutch speaking nursing units: For a fixed compensation the charge nurse, in collaboration with other involved nurses, wrote three patient cases themselves.

These were reviewed afterwards by the research team and fine tuned via mail correspondence.

In Dutch speaking hospitals three patient records were included and reviewed per participating nursing unit. In French speaking hospitals one or two patient records were reviewed per participating nursing unit. The rationale for this distinction is mentioned above in the description of the different data collection methods. To ensure maximal variability in staffing (and therefore cost) needs within the whole case collection, the charge nurse of each nursing unit was instructed to focus on a patient with a light, medium and heavy nursing care need, or light and heavy if only two cases were constructed. This needs estimation was based on a subjective impression of the charge nurse. Their experience in the context of the particular nursing unit was considered the best lead.

Appendix 4 en 5 represent the end result in terms of selected hospitals. The specialist typology of the participating nursing units per hospital is added. Representative sampling was the key criterion in this sub selection of hospitals and nursing units.

The participation in case construction consisted of 15 nursing units in 13 Dutch speaking hospitals and 46 nursing units in 22 French speaking hospitals. In total 38 general care, 6 paediatric, 6 geriatric and 7 intensive care units participated.

6.3.1.2 Case construction procedure

All data were gathered and all cases were written using a standard format (see Appendix 6). The NMDS-I and NMDS-II coding also took place in a consistent manner (see Appendix 7). To ensure a proper coverage of the workload range in nursing care, each participant was asked to select a light, medium and heavy nursing care patient case. This ranking was based on the participants own subjective judgment. Because the relative degree of workload is very context specific, a general definition of what constitutes the ranking levels was avoided. Workload variability in a specific context was assessed based on direct care experience. The ranking was discussed with the participants by the research team to ensure consistency.

Three forms of information are gathered: (1) information about the patient day of stay and care given, (2) information about the context of care and (3) registered NMDSI and NMDSII data.

1. Information about the patient day of stay and care given

The 24 hours of nursing care are described in detail separately for the morning, evening and night shift. Every shift consists of a different set of interventions which imply different staffing needs. A chronological description of all nursing interventions during the patient day of stay is constructed. Lab and other results of clinical investigations can be of importance. On an I.C.U. ward for example these can imply a substantive amount of additional nursing care.

2. Information about the context of care

A day of nurse caring has to be situated in the context of patient characteristics and the course of the whole hospital stay.

Patient characteristics such as age, gender, etc. are of influence on nursing care. The same holds true for example for the medication taken by the patient at home, which is continued during the hospital stay. The medical history of the patient gives additional information about co morbidities that necessitate additional care. Other important patient characteristics are also included in case construction.

Elements of the whole hospital stay include the reason of admission, the number of days of stay, the care rendered previously and the treatment stage which is now considered (e.g. first day post operative).

The above information is combined to constitute a picture of care, as completely as possible.

3. Registered NMDSI and NMDSII data

To analyse the relationship between cases and staffing needs NMDSII data are used. NMDSI data serve as a source of additional validation. Initially PRN was also considered as a source of validation. The necessary algorithms to code PRN weren't available, due to license protection reasons. Registration data are collected for each described day of nursing care.

6.3.1.3 Case sample description

A total of 141 cases are constructed (see Table 10). A sufficient reserve pool of cases has been developed. All cases have been translated to respectively French and Dutch.

Table 10: Case construction

Specialty typology	Total target	Constructed cases in Dutch	Constructed cases in French
Medical – surgical wards	40	27	31
Paediatric wards	20	24	5
ICU wards	20	24	5
Geriatric wards	20	20	5

The 5 paediatric, geriatric and intensive care cases based on patient data from French speaking hospitals were also used to assess case description differences between both regions. No relevant differences were identified. This ensures that all cases can be treated as a homogeneous group for inclusion in random distribution and rating.

Based on the quality of the cases the number of cases was further reduced for use in the actual rating procedure. Forty Medical-surgical cases, 25 Paediatric cases, 22 ICU cases and 25 Geriatric cases are withheld as source for rating staffing needs. The others were excluded based on clarity of description and content completeness. There is still a buffer compared to the total target set above.

6.3.2 Rating of patient cases

During the next study phase the patient cases were rated on nurse staffing needs. This process is described in the following subsections.

6.3.2.1 Rater selection and participation

During January and February 2006 all Belgian hospitals were invited to participate in the rating part of the study. For this aspect no difference was made between hospitals that participated in the former NMDS actualisation project or not. The goal was to create a broad a level of participation by the Belgian general hospitals as possible.

The following inclusion criteria are used to select potential candidates for rating participation:

- Raters are employed in the function of charge nurse or adjunct charge nurse at a paediatric (E), geriatric (G), intensive care (I), surgical (C) or general (D) ward in a Belgian general hospital.
- Raters have at least five years of clinical experience as a nurse.
- Raters have at least one year of experience as charge nurse or adjunct charge nurse.

A total of 227 nurses of 70 general hospitals applied as candidate raters. Twenty five of them aren't included, because the criteria weren't met or because the candidate had already

participated in previous rating rounds not part of the current study. An earlier rating experience could bias the current rating. Therefore these candidates are excluded.

So, 202 nurses of 69 general hospitals are withheld as raters. It concerns 76 nurses from the French speaking regions and 126 nurses from the Dutch speaking region.

The raters are divided in subgroups per specialty typology (Table 11).

Table 11: Raters by specialty typology (1)

Specialty typology	Total	Dutch speaking raters	French speaking raters
Medical – surgical wards	115	97	18
Paediatric wards	27	7	20
ICU wards	30	14	16
Geriatric wards	30	11	19

The division of raters over specialties was clearly unbalanced: CD group versus E, I and G. Two potential reasons can be cited: The CD group is an aggregate of multiple sub specialties such as cardiology, gastro enterology, nephrology, etc. As well in number of patients as in staffing proportion the CD group represents a larger part of a general hospital specialty mix. Therefore more raters belong to this subgroup (1). Previous rating studies, external to this research project, were merely focused on the E, I and G group. So the rater pool is already reduced compared to the CD group, since these nurses can't reapply (2). Ten ratings per case was set as a goal to obtain reliable data. In the current distribution this wasn't feasible. It was decided to redistribute the raters per specialty typology, as presented in Table 12.

Table 12: Raters by specialty typology (2)

Specialty typology	Total	Dutch speaking raters	French speaking raters
Medical – surgical wards	87	69	18
Paediatric wards	40	20	20
ICU wards	36	20	16
Geriatric wards	39	20	19

So, 28 Dutch speaking raters from the CD group were reassigned to the E, I and G groups. This reconfiguration made a goal of 10 ratings per case possible. Because this redistribution could lead to potential bias in ratings a safeguard was implemented. If a rater feels he of she had insufficient specialty experience, a consult with a colleague charge nurse or adjunct is indicated. Experience lacking concerning specific case elements can never be excluded. Therefore this safeguard was broadened to the whole of raters.

In the next step the cases were randomly assigned to the raters. The randomization procedure is fairly complex:

The cases were randomly assigned to the raters. Region participation wasn't used as a randomization criterion. The rating of staffing needs is considered on a national level. The assignment of cases should flow randomly over a regions level. This doesn't exclude the analysis of differences between regions. Sub grouping by specialty typology was the main stratification method to assign cases to raters. The following constraints had to be met:

- Goal = 10 ratings per case
- Restriction = minimum 8 cases per rater, maximum 12 cases per rater
- A case can't be rated by the same rater twice.

In cooperation with the Centre for Medical Biostatistics, K.U. Leuven, all cases could be randomly assigned to the raters per specialty subgroup, taking the constraints into account. The randomization output for the ICU subgroup is presented as an illustration in Appendix 8. Per rater the maximum twelve case numbers are listed in a random fashion. Each case number is presented at least ten times. All case numbers per rater are unique.

6.3.2.2 Rating procedure

All cases are rated during two consecutive rounds, as part of an adapted Delphi approach. I.e. all raters received feedback per case. The feedback enables the rater to change his or her original rating or not. The Delphi approach is a means to achieve greater consensus. In a traditional Delphi procedure, a final consensus meeting is also held. However, two objections can be stated in the current project framework. Firstly, it's on a practical level unfeasible to organize actual consensus meetings separately for each and every case, always with a changing rater composition. This would mean that each rater should attend 8 to 12 different meetings and that a total of 112 case meetings should be held. Secondly, the Delphi approach is used in this study as an aide for the rater. If he or she is unsure about one of the own ratings, feedback can give some clarification. Consensus is not the purpose in itself. If a nurse is convinced of a rating in contrast with the majority, conformity isn't a requisite. Accuracy of ratings is the main goal, based on the experience and expertise of caregivers.

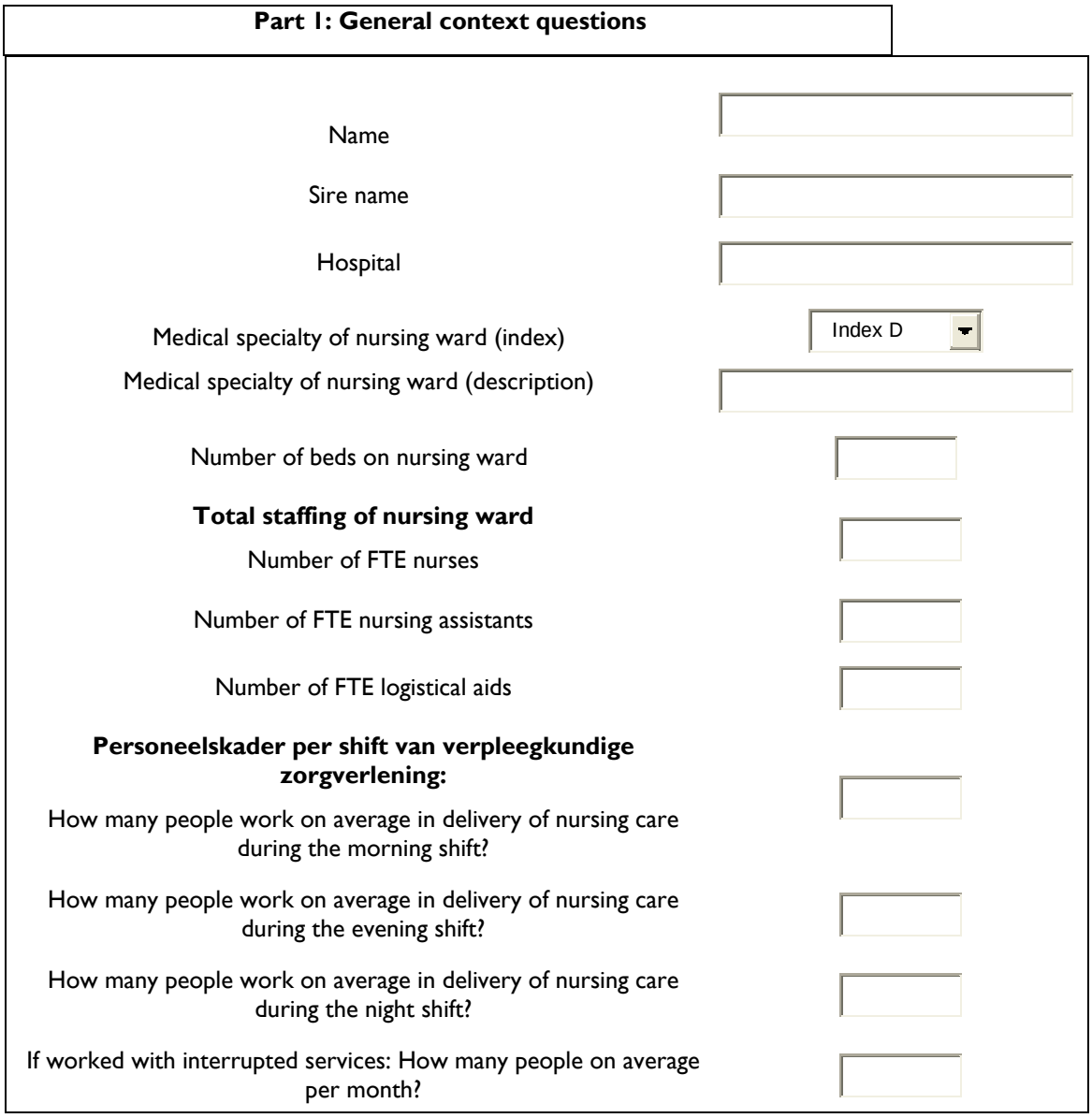
Cases are distributed to the raters on an individual basis by e-mail. The rating itself took place using a web based survey. Online each rater filled in a general information part once and a case specific part twice, during each round. The web application enabled the participants to send the data in parts, e.g. some case ratings on day 1, others on day 3, and so on. All data are collected in a tab delimited txt format. SAS® version 9.1 and Microsoft excel ® 2002 are used to clean and analyse the data. The rounds were subdivided in two groups based on specialty: CD round I and II, G E I round I and II. This enabled the beginning of the GEI ratings, before all CD cases were translated in both languages. For round I a minimal time frame of one month was provided; and for round II a minimum of two weeks. This provided the raters with sufficient time to review the cases and send their ratings. Reminder e-mails were sent to minimize non response. Additional individual rater phone calls were a further means of stimulation. The research team was always available by phone or e-mail for further information, questions or comments. These are gathered as an additional source of research data.

In the following subsection the content of the rating survey is described in detail.

6.3.2.3 Content of the rating survey

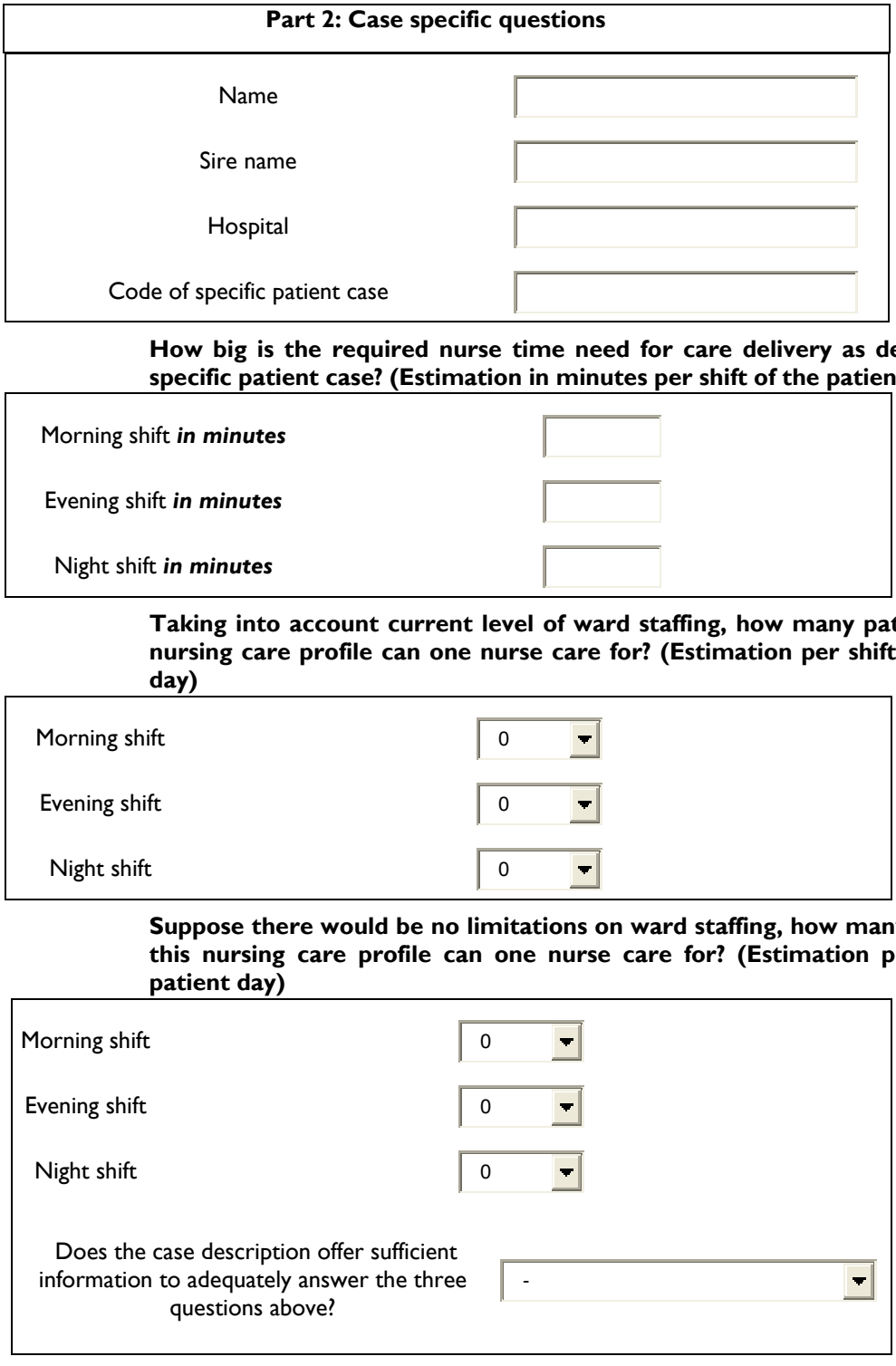
The following figure presents the main part of the rating survey in its authentic form:

Figure 12. Content of rating survey part I



The once deliverable part of the survey consisted of general information about the rater and the nursing ward he or she works as a frame of reference. Additional staffing information was also gathered. More specifically the rater name, hospital, clinical specialty by index and description, number of beds on the ward, number of FTE nurses as available staffing on the ward, number of FTE nursing aids as available staffing on the ward, number of logistics aids as available staffing on the ward, average number of working staff per shift of nursing care and number of staffing members that work within interrupted shifts, if these are present, during an average month are collected.

Figure 13. Content of rating survey part II



Part 2: Case specific questions

Name

Sire name

Hospital

Code of specific patient case

How big is the required nurse time need for care delivery as described in the specific patient case? (Estimation in minutes per shift of the patient day)

Morning shift ***in minutes***

Evening shift ***in minutes***

Night shift ***in minutes***

Taking into account current level of ward staffing, how many patients with this nursing care profile can one nurse care for? (Estimation per shift of the patient day)

Morning shift 0

Evening shift 0

Night shift 0

Suppose there would be no limitations on ward staffing, how many patients with this nursing care profile can one nurse care for? (Estimation per shift of the patient day)

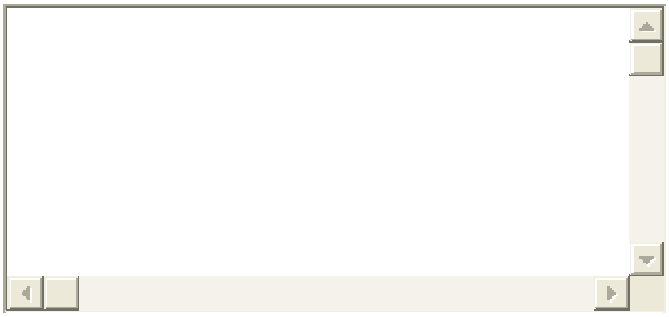
Morning shift 0

Evening shift 0

Night shift 0

Does the case description offer sufficient information to adequately answer the three questions above? -

Further comments or questions?



The case specific part of the survey consists of three main questions concerning staffing needs. Room for comments is provided and a rating of the case quality is asked. More specifically, the rater name, hospital, case code, needed staffing time in minutes for this patient profile per shift of nursing care, number of patients of this profile as feasible and quality assuring care load for one nurse per shift of nursing care (taking the current available staffing into account or not), case quality assessment, comments and questions are collected.

In addition to the three main rating questions concerning patient cases, a number of alternative staffing needs assessments were calculated. These included the following:

- The TISS patient classification system was used to review the staffing needs in intensive care patient cases. [121]
- The NARVEL patient classification system was used to review the staffing needs in pediatric care patient cases. [122]
- The SAN JOAQUIN patient classification system was used to review the staffing needs in surgical and general medicine care patient cases. [123]
- The AGGIR patient classification system was used to review the staffing needs in geriatric care patient cases. [124]

These case classifications were performed independently by two research participants, respectively of CZV, Catholic University Leuven, and CHU Liège.

6.4 STATISTICAL ANALYSIS

6.4.1 Aggregation from shift to patient day level

Before further analysis the rating data on a shift level had to be aggregated on a 24 hours level. This is a first step towards the patients whole stay level. Different scenarios were simulated: a separate shift approach, an unweighted summation approach and a summation based on a weighted mean approach.

The first gives insufficient information to address further research questions. The purpose is to assess staffing and financial needs on a 'whole patient' level. A 'day of hospital stay' forms the sublevel unit of analysis. These units are at the end combined into a unique sequential patient profile, capturing the global staffing needs with its inherent variation. Starting from staffing needs on a shift level adds sensitivity to staffing needs as an outcome measure.

Intervention demands at night are for example very different from the nursing workload during morning care. Also, the length of morning, evening and night shift differ daily within a fixed pattern. We suppose a 7/7/10 hour division. Therefore this difference has to be taken into account.

So a prolonged separate shift approach is unusable and an unweighted summation approach lacks sensitivity. The third approach, a weighted mean summation, combines all the relevant

information. The 24 hours measure is equal to three times the weighted shift rating. The 'estimated nursing time in minutes', the first question variable, is treated in this way.

The variables 'number of patients within current staffing availability' and 'number of patients without staffing limitation', defined by question two and three, are not continuous in nature. A patient can't be divided in performing formulary arithmetic's. In addition, a summation of patients over three shifts is unwarranted. The same patient set from the morning is treated during the following shifts. Adding them as new patients each time would imply a three times overestimation. So for question two and three the weighted median without summation is considered as the relevant outcome measure. The weighting still captures differences in nursing workload by shift.

6.4.2 Robust selection of 'staffing needs'

Statistical analysis of the three rating questions on a patient day of stay level started with central and distribution descriptive measures by case. These are illustrated below by an example. The nature of the distribution in terms of normality was also checked. The intensive care, paediatric care and geriatric care descriptive measures show similar characteristics as explained below. Comprehensive data about all cases can be obtained upon request. In Appendix 9 more comprehensive tables and figures can be found for purposes of overview.

Regular frequency analysis pointed to a lack of a sufficient number of raters in case PE07. Four ratings falls below the chosen minimal level of five, so this case is excluded from further analysis.

The analysis is explained below step by step using the following legend:

- 'weighted_min' = the number of minutes of required nursing care time to deliver adequate care for patient X, as described in patient case X during 24 hours, taking into account shift differences.
- 'weighted_pat' = the number of patients identical to patient X, as described in patient case X for which one nurse can deliver adequate care during 24 hours, taking into account shift differences. Regular staff limits are imposed.
- 'weighted_optpat' = the number of patients identical to patient X, as described in patient case X for which one nurse can deliver adequate care during 24 hours, taking into account shift differences. No staff limits are imposed.

Table 13: Descriptive measures of a random patient case

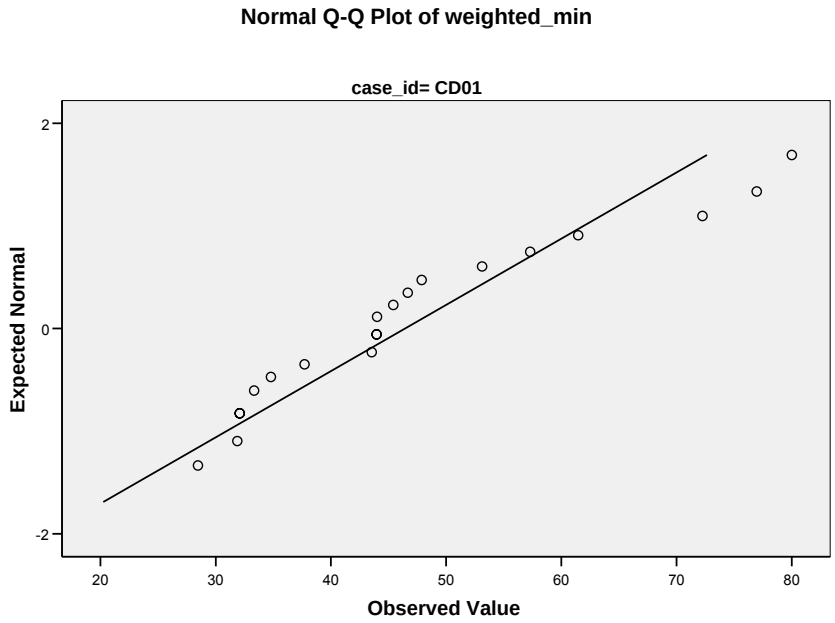
			Statistic	Std. Error
weighted_min	Mean		46,4385	3,38040
	95% Confidence Interval for Mean	Lower Bound	39,3871	
		Upper Bound	53,4899	
	5% Trimmed Mean		45,5853	
	Median		43,9583	
weighted_pat	Mean		7,6349	0,69378
	95% Confidence Interval for Mean	Lower Bound	6,1877	
		Upper Bound	9,0821	
	5% Trimmed Mean		7,7350	
	Median		8,2083	
weighted_optpat	Mean		6,3968	0,57414
	95% Confidence Interval for Mean	Lower Bound	5,1992	
		Upper Bound	7,5945	
	5% Trimmed Mean		6,5187	
	Median		6,3333	

During each shift patient X requires 46 minutes of care, estimated by average. The trimmed mean and median of 'weighted_min' differ slightly from this value due to a right skewed distribution and outlier influence. The rating of nursing time needs varies from a minimum of 28.38 minutes up to 80 minutes. There is a considerate degree of variability present between ratings. On a case level, interrater reliability is low. This applies to most patient cases, even after using a Delphi approach.

Table 14: Normality measures for randomly selected example patient case rating

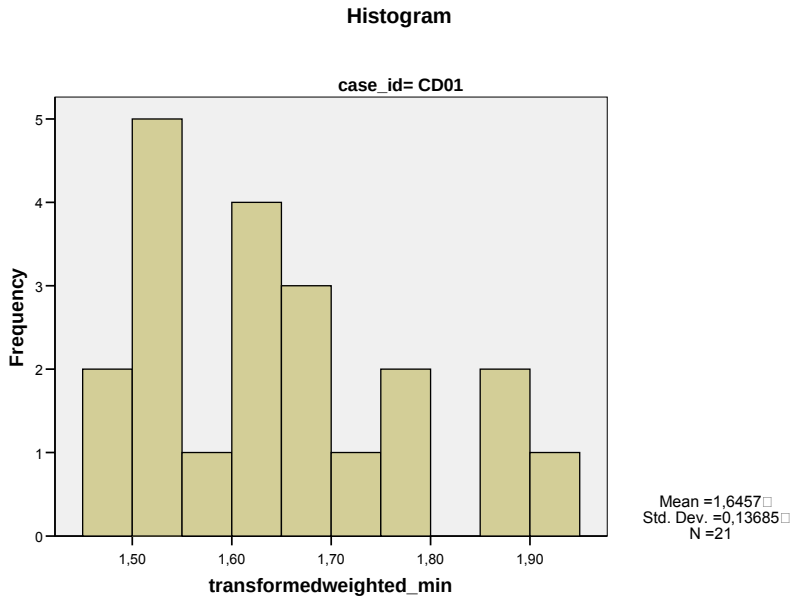
	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
weighted_min	0,177	21	0,083	0,892	21	0,024
weighted_pat	0,123	21	,200	0,967	21	0,668
weighted_optpat	0,129	21	,200	0,965	21	0,631

Figure 14: Normality figure for a randomly selected example patient case rating



As in this case, 'weighted_min' sometimes doesn't pass standard normality tests. Therefore a transformation scenario using log 10 was tested. As you can see in the histogram below, this improved normality only slightly. Using transformed 'weighted min' as a starting point to append all further analyses on would complicate clarity and interpretations. Because of this reason the transformation option on a case level was dismissed.

Figure 15: Histogram of example patient case rating after log10 transformation



To account for skew ness and outlier influence another avenue was pursued. Different alternative central measures taking these features into account were examined. For 'weighted_min', next to a trimmed mean, standard M – estimators were considered (see table below illustrating our random case).

Table 15: M – estimators of example patient case rating

	Huber's M-Estimator	Tukey's Biweight
weighted_min	43,7890	43,3050
weighted_pat	7,9118	8,0321
weighted_optpat	6,6210	6,7822

The Shapiro-Wilk normality test remains robust, even in small samples, contrary to the Kolmogorov-Smirnov test. If the Shapiro-Wilk normality test showed a significant deviation from normality ($p < 0.05$), the mean was considered a fair measure. Else the Huber robust mean was selected. A trimmed mean is inappropriate, because an asymmetrical distribution to the right remains too far from the arithmetic mean. The choice of Huber compared to other M – estimators is rather strategic. Huber gives a higher output. In terms of staffing and costing it is therefore preferred. This approach avoids a systematic underestimation. For the illustrating patient case 43.7890 was chosen as an adequate central measure of 'weighted_min'.

For a given variable, determining the best way to describe what's average depends upon:

- The scale of measurement of the variable
- The shape of the distribution
- Presence of outliers

While the mean and the sample variance represent the best estimates when the sample comes from a normal population, they can be greatly affected by the presence of unusual or extreme values and is pulled in the direction of the skew of the distribution. These *outliers* are sample values that cause surprise in relation to the majority of the sample. This is not a pejorative term; outliers may be correct, but to regard as extreme, atypical values. These extreme values can play havoc with standard statistical methods, and many *robust* and *resistant* methods have been developed since 1960 to be less sensitive to outliers.

On the other hand, the median is insensitive to outliers, addition or removal of extreme values has little effect on it. The median is called a resistant measure. Although the median is an intuitive, simple measure of location, there are better estimators of location if we are willing to make some assumptions about the population from which our data originate. Estimators that depend on simple, fairly non-restrictive assumption about the underlying distribution and are not sensitive to the assumptions are called robust estimators.

Robust methods resist outlier influence by down weighting values the further they are from the centre of the sample. Simultaneously, a good robust method is also quite efficient, though not optimal, in the ideal case of data coming from a Gaussian population. They are also efficient in other cases as well, e.g., those in which the underlying distribution is heavy-tailed.

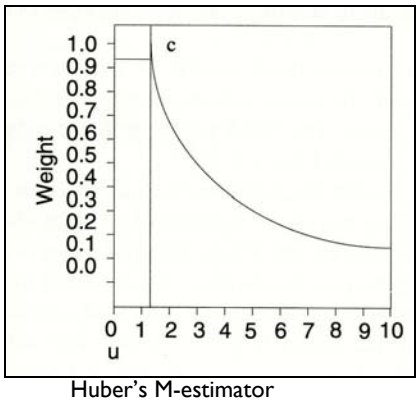
Truncation methods make an ad hoc assumption about the percent of outliers in the data set. A simple example would truncate, or trim, the 5% smallest and 5% largest data values. In this case, the lower and upper 5% of the data would get weight zero, the middle 90% would then be used. If N is large, the resulting trimmed mean will be a more stable estimate of central tendency than the arithmetic mean. On the other hand, if N is small, this may not be wise since it might jeopardize the external validity of the sample.

Huber's (1964) M-estimators represent a very flexible and general class of estimators which played an important role in the development of robust statistics and in the construction of robust procedures and represents also an interesting alternative at truncation methods. *M*-estimators are based on the approach used in maximum likelihood estimators.

The Huber influence function decreases in an asymptotic way the influence of outliers. This function is given by:

$$\psi(u) = \begin{cases} u, & \text{si } |u| \leq a \\ a\frac{u}{|u|}, & \text{si } |u| > a \end{cases}$$

The next figure shows for different values of a standardized distance from the estimate of location, the evolution of weight value.

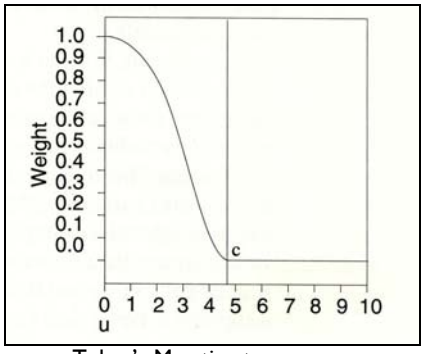


Huber's M-estimator

The Tukey influence function rejects completely outliers and gives them a null weight. This function is given by:

$$\psi(u) = \begin{cases} u(C^2 - u^2)^2, & \text{si } |u| \leq C \\ 0 & \text{, sinon} \end{cases}$$

The next figure shows for different values of a standardized distance from the estimate of location, the evolution of weight value.



Tukey's M-estimator

In our study, we chose Huber's estimator rather than Tukey because it was not natural that actual values are rejected.

The robust selected 'weighted_min' per case is used in the further model construction. It is a clearer cut outcome than 'weighted_pat' and 'weighted_optpat'. Free comments of rater analysis confirmed that this rate was most easy to interpret and estimate in practice.

6.5 RESULTS

6.5.1 Results from comments and non response

During the course of the ratings the main concern reported in the 'free commentary' part of the survey was a too narrow definition of rating categories for the second and third question. These question the number of patients per nurse, if the case staffing needs are extrapolated to a nursing ward level. Because it wasn't intended to impose an upper level, the opportunity was provided to go beyond the predefined categories. This possibility was clearly communicated to all raters.

The raters reported that case rating is a feasible approach, but reviewing and assessing 8 to 12 patient cases takes sometimes several hours of time. A sample of non responders was questioned by phone. The workload and competing priorities in time management were most frequently stated as reasons. Otherwise non response seemed distributed randomly over participating general hospitals.

The response during the whole rating process equals to 92% of expected ratings.

6.5.2 Internal consistency of ratings

Correlations below illustrate a high level of internal consistency over all cases between the three rating questions. There is a strong relationship between the estimated required nursing time per patient and the estimated appropriate number of those patients in care delivery per nurse. The relation between 'weighted_pat' and 'weighted_optpat' is almost perfect. It seems as if the presence of staffing limitations doesn't affect the number of patients per nurse. This is also quite logical since the question concerns one nurse, independently of the complementary number of staffing. Caution in interference is however indicated. The free commentary and the individual ratings data indicate some confusion about the staffing limitation effect. Some raters applied consistently no relationship, others a positive relationship and a third group a negative relationship.

So there is sufficient evidence for a strong internal consistency. Staffing limit effects are insufficiently clear.

Table 16: Internal consistency correlation matrix

		weighted_min	weighted_pat	weighted_optpat
weighted_min	Pearson Correlation	1	-,763(**)	-,784(**)
	Sig. (2-tailed)		0,000	0,000
	N	112	112	112
weighted_pat	Pearson Correlation	-,763(**)	1	,970(**)
	Sig. (2-tailed)	0,000		0,000
	N	112	112	112
weighted_optpat	Pearson Correlation	-,784(**)	,970(**)	1
	Sig. (2-tailed)	0,000	0,000	
	N	112	112	112

**. Correlation is significant at the 0.01 level (2-tailed).

In a similar way internal consistency is confirmed by nonparametric correlations such as Spearman's Rho and Kendall Tau. Correlations only differ slightly, with an almost negligible deviation. The same is true for a separate analysis of internal consistency on a shift level. All correlations are significant on a $p < 0.01$ level.

6.5.3 Concurrent validity

Next to internal consistency, concurrent validity was checked on a case level using TISS, AGGIR, NARVEL and San Joaquin.

Table 17: External consistency correlation matrix

			Robust selection weighted min
Kendall's Tau	**TISSTOT**	Correlation Coefficient	,741(**)
		Sig. (2-tailed)	0,000
		N	22
	SAN JOAQUIN	Correlation Coefficient	,272(*)
		Sig. (2-tailed)	0,036
		N	40
	NARVEL	Correlation Coefficient	0,142
		Sig. (2-tailed)	0,351
		N	25
	AGGIR	Correlation Coefficient	-,405(**)
		Sig. (2-tailed)	0,010
		N	25

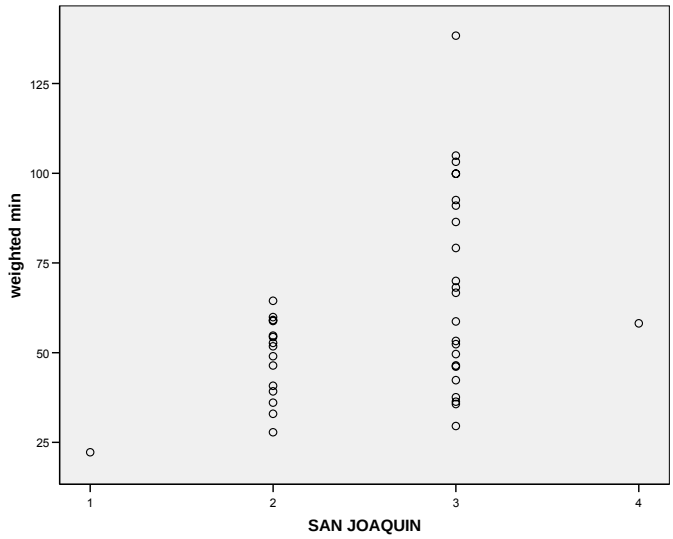
** Correlation is significant at the 0.01 level (2-tailed).

* Correlation is significant at the 0.05 level (2-tailed).

A very strong relationship exists between the estimated time and the TISS patient classification system regarding intensive care. This confirms previous research.

The geriatric AGGIR – estimated time relationship is also strong. However the relationship with San Joaquin for general care is weak and with NARVEL for paediatric care is the relationship non existing. NARVEL is in general considered to be an inadequate patient classification system, so absence of relationship in this area is not surprising. NARVEL was only used, because of the lack of other known appropriate paediatric patient classification systems. Research with regard to the relation with other existing paediatric patient classification systems is therefore recommended. The relationship of estimated time in general patient cases with San Joaquin is a bit obscured by the CD case coverage of nursing workload towards both extremes. As graphically presented in the figure below, only one CD case falls within SJ class 1, one within class 4 and the majority within class 2 and 3. This is also due to the very high criteria in SJ which define class 4, e.g. constant monitoring and observation of patient's medical condition. At the other end of the continuum patients within class 1 are almost totally independent for activities of daily living. Both situations are rare on a surgical or general care ward. Hence the lack of sufficient variability of the patient case nursing time needs within the SJ classification. This leads to insufficient evidence of appropriate external validity concerning general care.

Figure 16: General care patient case distribution within the San Joaquin classification



6.5.4 Consistency across regions

In a following step a potential difference in rating based on region was investigated. Differences in healthcare and culture between the French speaking region and the Dutch speaking region are known. This could affect the rating of estimated appropriate nursing time per patient case. This hypothesis was assessed using a Forest plot methodology. The plot is presented in the figure below (see next page).

Review: regionale vergelijking
Comparison: 01 frans versus nederlands
Outcome: 01 min rating zorgbehoefte, gewogen 24u

Study or sub-category	N	French region Mean (SD)	N	Dutch region Mean (SD)	SMD (fixed) 95% CI	Weight %	SMD (fixed) 95% CI
CD01	6	61.33(18.95)	16	40.48(9.02)		1.56	1.63 [0.55, 2.71]
CD02	4	42.53(13.26)	12	34.29(17.61)		1.38	0.46 [-0.68, 1.61]
CD03	2	91.10(6.98)	4	28.86(3.38)		0.02	10.93 [0.24, 21.62]
CD04	4	47.85(7.39)	8	30.16(16.93)		1.05	1.11 [-0.21, 2.43]
CD05	3	92.14(6.05)	4	34.32(13.43)		0.13	4.39 [0.60, 8.18]
CD06	2	81.25(62.87)	9	53.05(12.04)		0.68	1.08 [-0.55, 2.71]
CD07	4	34.70(15.14)	10	32.26(5.69)		1.34	0.25 [-0.91, 1.42]
CD08	8	64.05(27.47)	17	46.88(17.34)		2.38	0.79 [-0.08, 1.67]
CD09	7	75.27(25.23)	9	65.85(28.48)		1.83	0.33 [-0.67, 1.32]
CD10	12	31.51(13.86)	5	18.91(0.97)		1.47	1.01 [-0.11, 2.12]
CD11	4	71.88(28.64)	8	41.71(15.03)		0.96	1.39 [0.01, 2.76]
CD12	2	86.45(38.59)	9	52.54(18.67)		0.63	1.42 [-0.28, 3.12]
CD13	5	75.35(30.98)	12	44.06(13.39)		1.28	1.51 [0.32, 2.70]
CD14	4	60.51(33.13)	7	51.47(30.27)		1.19	0.26 [-0.97, 1.50]
CD15	5	126.80(40.73)	7	80.64(32.13)		1.10	1.19 [-0.10, 2.48]
CD16	2	78.97(31.02)	7	112.32(40.07)		0.68	-0.76 [-2.40, 0.88]
CD17	5	61.35(20.39)	9	38.13(13.60)		1.18	1.34 [0.10, 2.58]
CD18	3	71.61(46.34)	9	41.49(12.48)		0.89	1.18 [-0.25, 2.61]
CD19	6	49.54(7.49)	6	32.01(11.10)		0.92	1.71 [0.30, 3.11]
CD20	2	83.02(1.62)	7	59.15(30.38)		0.68	0.75 [-0.88, 2.39]
CD21	6	62.39(18.74)	11	47.45(22.43)		1.72	0.67 [-0.36, 1.69]
CD22	4	49.08(20.44)	7	38.47(11.44)		1.12	0.64 [-0.63, 1.92]
CD23	2	36.48(22.73)	9	58.32(22.54)		0.71	-0.89 [-2.49, 0.72]
CD24	4	55.65(14.51)	8	30.98(15.68)		0.92	1.48 [0.08, 2.89]
CD25	4	124.97(54.95)	6	88.68(39.64)		1.03	0.71 [-0.61, 2.04]
CD26	5	113.63(93.55)	8	69.42(17.55)		1.34	0.71 [-0.46, 1.87]
CD27	5	91.75(29.44)	8	51.00(16.84)		0.98	1.70 [0.34, 3.07]
CD28	3	61.69(13.02)	7	39.43(19.43)		0.82	1.11 [-0.38, 2.61]
CD29	5	54.30(29.69)	6	23.68(5.98)		0.94	1.38 [-0.01, 2.77]
CD30	5	94.30(28.57)	9	89.11(28.53)		1.51	0.17 [-0.93, 1.27]
CD31	7	82.11(18.31)	11	44.01(44.42)		1.76	0.98 [-0.03, 2.00]
CD32	7	130.83(74.25)	4	151.51(24.01)		1.19	-0.30 [-1.54, 0.93]
CD33	2	98.04(7.42)	10	52.30(28.66)		0.63	1.55 [-0.15, 3.24]
CD34	5	42.98(15.04)	11	32.37(7.70)		1.43	0.97 [-0.16, 2.10]
CD35	6	81.41(22.26)	7	77.25(25.00)		1.52	0.16 [-0.93, 1.26]
CD36	4	28.62(19.16)	10	19.68(6.51)		1.25	0.75 [-0.45, 1.96]
CD37	6	32.04(12.38)	11	28.20(9.80)		1.81	0.34 [-0.66, 1.34]
CD38	2	88.79(9.66)	6	104.30(27.95)		0.68	-0.52 [-2.16, 1.12]
CD39	5	94.13(46.69)	10	53.73(15.01)		1.25	1.32 [0.12, 2.53]
CD40	7	124.79(56.95)	10	69.97(19.23)		1.52	1.33 [0.24, 2.43]
G01	9	71.34(13.23)	10	48.93(20.81)		1.82	1.21 [0.21, 2.21]
G02	6	40.04(22.37)	8	24.49(11.20)		1.44	0.87 [-0.26, 1.99]
G03	3	32.65(15.46)	7	15.64(3.86)		0.63	1.82 [0.13, 3.52]
G04	4	52.93(10.45)	6	46.84(17.00)		1.11	0.37 [-0.91, 1.65]
G05	3	69.43(10.62)	5	49.75(23.94)		0.76	0.84 [-0.71, 2.38]
G06	3	22.40(7.98)	7	14.33(4.92)		0.78	1.25 [-0.28, 2.77]
G07	3	16.13(5.46)	7	13.95(5.69)		0.97	0.35 [-1.02, 1.72]
G08	4	55.33(13.49)	6	42.01(12.06)		0.96	0.95 [-0.42, 2.33]
G09	5	57.86(18.24)	4	42.37(15.43)		0.92	0.81 [-0.60, 2.21]
G10	9	62.88(23.20)	10	44.53(17.98)		2.01	0.85 [-0.10, 1.80]
G11	7	44.27(11.08)	7	40.89(22.43)		1.65	0.18 [-0.87, 1.23]
G12	2	79.16(8.24)	9	45.37(18.03)		0.56	1.79 [0.00, 3.59]
G13	7	56.83(16.14)	8	53.78(25.50)		1.76	0.13 [-0.88, 1.15]
G14	11	40.76(18.24)	6	36.37(13.28)		1.82	0.25 [-0.75, 1.25]
G15	8	39.27(13.94)	6	31.59(15.17)		1.56	0.50 [-0.58, 1.58]
G16	5	94.67(46.59)	5	68.10(26.65)		1.09	0.63 [-0.66, 1.92]
G17	9	63.65(38.80)	8	39.70(15.22)		1.84	0.75 [-0.24, 1.75]
G18	9	45.85(16.38)	5	37.92(9.95)		1.46	0.51 [-0.61, 1.63]
G19	5	51.92(12.73)	5	45.23(12.82)		1.13	0.47 [-0.79, 1.74]
G20	4	49.05(18.40)	9	26.28(7.25)		0.86	1.85 [0.40, 3.31]
G21	7	69.72(30.23)	6	55.56(14.20)		1.45	0.54 [-0.58, 1.66]
G22	6	56.74(25.94)	7	49.09(12.48)		1.49	0.36 [-0.74, 1.46]
G23	4	33.12(12.71)	5	37.69(11.08)		1.03	-0.34 [-1.68, 0.99]
G24	6	62.88(29.63)	6	44.61(14.70)		1.29	0.72 [-0.46, 1.91]
G25	8	55.73(16.19)	8	45.22(13.17)		1.76	0.67 [-0.34, 1.69]
IC02	1	317.54(0.00)	6	382.61(76.68)			Not estimable
IC04	2	431.61(122.55)	4	211.17(22.22)		0.18	2.75 [-0.40, 5.89]
IC05	2	302.86(266.67)	4	258.73(42.90)		0.62	0.26 [-1.46, 1.97]
IC06	2	162.95(15.08)	7	245.38(23.01)		0.27	-3.32 [-5.90, -0.74]
IC07	3	301.05(261.65)	8	271.81(70.48)		1.03	0.19 [-1.14, 1.52]
IC09	2	340.72(168.84)	4	250.40(54.30)		0.53	0.75 [-1.10, 2.59]
IC15	1	169.92(0.00)	5	230.86(20.00)			Not estimable
IC16	3	121.47(70.01)	3	148.68(34.89)		0.67	-0.39 [-2.04, 1.25]
IC17	4	220.62(138.59)	4	154.05(21.85)		0.87	0.58 [-0.86, 2.03]
IC18	1	82.21(0.00)	5	124.37(27.52)			Not estimable
IC19	2	60.54(8.95)	7	145.57(41.85)		0.47	-1.94 [-3.92, 0.03]
IC20	3	106.05(69.82)	3	73.94(33.80)		0.66	0.47 [-1.19, 2.13]
IC22	1	470.89(0.00)	6	332.17(112.63)			Not estimable
IC23	2	381.87(125.69)	3	309.38(17.94)		0.44	0.71 [-1.32, 2.74]
IC24	4	325.00(170.72)	6	306.02(74.37)		1.13	0.14 [-1.12, 1.41]
PE01	4	94.78(62.00)	2	132.29(35.06)		0.58	-0.53 [-2.30, 1.24]
PE02	4	50.79(24.24)	2	53.64(14.87)		0.63	-0.10 [-1.80, 1.60]
PE03	2	31.16(6.36)	6	57.97(13.80)		0.44	-1.81 [-3.84, 0.22]
PE04	5	69.67(51.42)	3	64.86(24.18)		0.89	0.09 [-1.34, 1.53]
PE05	3	77.59(19.22)	5	62.74(31.75)		0.85	0.46 [-1.01, 1.92]
PE06	2	87.60(5.45)	3	73.40(11.82)		0.36	1.02 [-1.24, 3.27]
PE07	1	22.92(0.00)	3	32.93(3.70)			Not estimable
PE08	6	181.70(85.38)	5	153.40(54.25)		1.26	0.35 [-0.85, 1.55]
PE09	4	103.39(105.16)	1	56.88(0.00)			Not estimable
PE10	6	70.95(30.30)	3	77.50(15.19)		0.94	-0.22 [-1.61, 1.18]
PE11	2	80.00(8.83)	4	93.34(32.96)		0.60	-0.37 [-2.10, 1.36]
PE12	1	156.96(0.00)	6	119.13(61.91)			Not estimable
PE13	1	137.50(0.00)	5	90.12(20.47)			Not estimable
PE14	2	168.02(97.96)	5	137.14(22.02)		0.63	0.54 [-1.15, 2.24]
PE15	1	28.75(0.00)	4	89.14(13.73)			Not estimable
PE16	2	78.93(40.98)	5	85.04(16.11)		0.67	-0.22 [-1.87, 1.43]
PE17	2	105.83(68.94)	9	93.77(27.55)		0.77	0.32 [-1.22, 1.86]
PE18	3	71.00(58.72)	5	92.48(73.00)		0.87	-0.27 [-1.72, 1.17]
PE19	3	91.83(14.75)	5	106.79(46.45)		0.87	-0.33 [-1.78, 1.12]
PE20	5	122.45(76.50)	5	86.99(29.37)		1.11	0.55 [-0.73, 1.83]
PE21	2	114.14(31.73)	4	136.51(67.16)		0.61	-0.30 [-2.02, 1.42]
PE22	3	217.06(130.37)	7	139.70(57.68)		0.88	0.85 [-0.58, 2.29]
PE23	2	143.75(28.87)	6	105.98(85.04)		0.69	0.42 [-1.21, 2.04]
PE24	4	102.75(35.61)	3	42.84(16.93)		0.45	1.71 [-0.31, 3.72]
PE25	5	115.40(44.32)	7	128.12(55.16)		1.37	-0.23 [-1.38, 0.92]
Total (95% CI)	439		702			100.00	0.61 [0.47, 0.74]

Test for heterogeneity: Chi² = 98.60, df = 95 (P = 0.38), I² = 3.6%
Test for overall effect: Z = 8.80 (P < 0.00001)

The forest plot analyses the French – Dutch ratings taking into account the difference in average and variability of estimated time with patient cases as units of analysis. For some cases this approach was unfeasible, because one of the compared regions comprised only one rater. The majority of cases however indicate a consistent difference towards higher ratings in the French speaking region. The overall effect is highly significant: $p < 0.00001$.

6.6 DISCUSSION AND CONCLUSIONS

The methods of case construction and rating have proven to be a good way to assess justified nurse staffing needs in daily patient care situations. It provides a rich framework to capture the complexity and variability which are inherent to nursing care. In addition, the rating ensures a broad representation of clinical practice in a variety of general hospitals. Clinical expertise and experience are key drivers in determining justified staffing needs instead of actual staffing levels. However, the used data collection methods also have some disadvantages. Cases should be constructed in a sufficiently high number and rated by a sufficiently high number of raters. Both conditions were fulfilled in the present study, but it consumed a high amount of time and other resources. Constructing a comprehensive patient case takes time. The rating requires sufficient follow-up and coordination. A web based application was one of the factors that made this possible. The Delphi part in the methodology didn't add much value to the original data. Not all Delphi aspects can be implemented, e.g. consensus meetings, due to feasibility reasons. Interrater reliability was only partially improved.

Validity and reliability are both very striking as results of this study. Focusing on a separate patient case, as a building block in relative staffing needs assessment, variability is high and interrater reliability is low. The same variability in estimates is found across regions. However, consistency and criterion validity measures show strong to exceptionally strong positive results. This is true between alternative patient case inquiries as measures of internal consistency; and using external comparisons such as TISS and AGGIR.

The goal of this study is to construct a model for an appropriate relative division of resources within the limits of a certain fixed budget. This implies that a difference in absolute terms between the estimated appropriate nursing time and other similar measures such as patient classification outcomes, is of little importance. As long as mutual trends are similar per assessed patient day of stay, the resource output will also be divided consistently.

Same is true for regional differences. Estimated appropriate nursing time as a measure on a national level diffuses the difference in rating between both regions over the whole. This outcome can be considered as a middle way taking into account both regional estimations. The end result will be that hospitals in the north will be attributed a bit more staffing and nursing financing than based on a unique regional needs estimation approach. Hospitals in the south will receive a little less than based on their own regionally estimated nursing time. In this way differences are levelled out in division of resources. Further research is indicated to study if the difference in estimation holds ground in actual nursing care differences. Other conclusions can be considered premature.

7 NURSE STAFFING NEEDS IN A GENERAL HOSPITAL ON A NURSING INTERVENTION LEVEL

7.1 INTRODUCTION

One of the oldest nursing activity studies was published in 1934 as an attempt to answer the question: what is good nursing? John and Pfefferkorn (1934) assembled a list of 798 nursing activities [114]. Activity methods involve an assessment of the patient for the activities involved in providing the nursing care required. Each activity is 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 [116]. Some reported difficulties are: perceptual differences of what constitutes an activity [115] and difficulties to assess time requirements for non physical needs of patients [116].

Hypothesis is that the summation of separate intervention nursing time estimated needs will be consistently higher than estimated time for a patient case as a whole. This is mainly due to the interactions between nursing interventions in reality and the competence to do multiple tasks simultaneously, as explained in the previous chapter. For example hygienic care, education and emotional assistance often interact with each other and can't be separated in time from each other. If one adds them independently an overestimation is likely to occur.

Otherwise, estimating the appropriate total nursing time for a 24 hours patient case, in this study subdivided by shift, is also susceptible to bias. It is easy to imagine that certain interventions dominate others in the mind of a rater. These other, maybe shorter, but also important interventions could be discounted in composing the total picture. So an underestimation is likely to occur. Combining both forms of bias, one expects the real nursing staffing need to be situated between both perceptions. The relation between both methods to assess nurse time needs is the subject of the next chapter. Firstly, this chapter explains the activity method used in the current study and its resulting findings. Since NMDSII can be considered a thorough overview of nursing interventions based on international grounding, nurse practice participation and statistical analysis, it is expected to be a useful tool to apply an activity method in staffing issues.

7.2 RESEARCH QUESTION

This study part is aimed at the following research questions:

- Is it possible to make a transfer from actual towards justified nurse staffing needs taking into account nursing care requirements to ensure quality of care?
- Which are the estimated nursing time needs with regard to each of the nursing interventions as registered in NMDSII to ensure quality of care?

7.3 METHODOLOGY

7.3.1 Rater selection and participation

All NMDSII interventions were rated separately, independently of any patient case, by 20 raters. These raters were selected randomly from the total rater pool, irrespective of specialty subgroup or region participation, to rate staffing needs on a nursing intervention level.

7.3.2 Rating procedure

The rating was based on a survey distributed by e-mail. The survey is aimed at the following question: How much time does a nursing team spend on average on… [specific NMDSII nursing intervention]…in caring for a typical patient to ensure quality of care? This question is posed for each of the 79 NMDSII items and is further subdivided by NMDSII sub item categories. To account for the potential problem as mentioned above concerning what constitutes a nursing intervention, the main question is for each NMDSII item embedded into the official NMDSII registration manual as available in August 2006. This manual provides a clear definition incorporating all relevant modalities of execution. An example, the item A100: Structured physical exercises (Dutch version) is presented below. It first defines the content of physical exercises. Then it makes a differentiation based on different patient needs concerning physical exercises, in this example passive versus active exercises. The nursing time assessment question is subdivided for each of the different intervention or patient modalities. NMDSII can in fact be considered as a very thorough overview of nursing interventions in which most items already incorporate typical nominal, ordinal or continual primary elements as present in typical patient classification systems. So NMDSII is expected to be sensitive in taking patient factors into account.

Example of survey content

A100: Gestructureerde lichamelijke oefeningen

Definitie

Passieve of actieve lichamelijke oefeningen, begeleid en opgevolgd door een zorgverlener, die geïntegreerd zijn in een standaard revalidatieplan of specifiek voorgeschreven zijn voor een patiënt. Deze oefen- of revalidatieplannen zijn uitgeschreven door een kinesist, arts, verpleegkundige of een multidisciplinair team. In het patiëntendossier is het oefen- of revalidatieplan aanwezig.

Passief: bewegingsoefeningen van de patiënt die uitgevoerd worden door een zorgverlener.

Actief: de zorgverlener *begeleidt* EN *volgt* de oefeningen op die uitgevoerd worden door de patiënt (bv. de zorgverlener verzekert de continuering van een stapoefening die opgesteld werd door de kinesist). De zorgverlener is tijdens deze oefeningen *permanent* bij de patiënt aanwezig.

Vraagstelling

Hoeveel tijd besteedt de verpleegkundige equipe **gedurende één verzorgingdag (24u)** gemiddeld aan de uitvoering van **passieve oefeningen** bij een doorsnee patiënt met geldige indicatie, om kwaliteitsvolle zorg te verzekeren? (minuten)

Hoeveel tijd besteedt de verpleegkundige equipe gemiddeld **gedurende één verzorgingdag (24u)** aan de uitvoering van **actieve oefeningen** bij een doorsnee patiënt met geldige indicatie, om kwaliteitsvolle zorg te verzekeren? (minuten)

It is important to bear in mind that the survey of intervention times utilized a double approach. Most interventions could be assessed directly on a 24 hours basis by the 20 respondents. For example: how much nursing time is needed for adequate hygienic care with partial assistance can be estimated for 24 hours. But e.g. blood administration depends on the frequency of administrations and the number of units per administration. Therefore this item is assessed on a blood unit level. The following items are frequency based: B300 (bladder catheterisation), C300 (patient transport), H100-500 (medication items), L200-500 (wound care), N100-400 (blood administration, blood sampling, and care for artificial entry points), V200 (pressure ulcer preventive repositioning), V500 (non blood sampling), W200 (ante partum care), W400 (post partum care) and Z200 (medical procedure support). The other interventions are estimated directly on a 24 hours basis.

This activity method takes patient factors and other direct care differences into account. However, an explicit distinction between front office and back office was made. Front office includes all nursing care during patient contacts. Back office consists of indirect supporting

nursing tasks such as nursing care administration, preventive hand hygiene between nursing care episodes, nursing care material handling, etc. Back office aspects of nursing care are not included in the survey. This exclusion was made because of a principal point of view. Nursing time is a scarce product, which should primarily be divided based on direct patient needs and the degree to which they differ. In contrast, most indirect back office nursing tasks are generic in nature. It means that administration, material handling, providing linen, answering the phone, etc. are very difficult to allocate to individual patients with individual needs and in surplus: on average an equal rate of consumption of these indirect resources can be expected between patients. A good patient record should be kept for every patient with more or less equal time consumption. Since the number of patients is already accounted for in another way in nursing financing allocation, to include it as an additional driver makes no sense. Other back office elements that do differ between patients are by definition driven by direct care elements that are already included in the survey. E.g. a patient with five instead of 1 intermittent bladder catheterizations has a greater need for collection of bladder catheterization materials. But the frequency of bladder catheterization is already included in NMDSII and the activity method survey.

The care is considered as delivered by a whole nursing team, irrespective of a shift of care approach. A questioning of three shifts per intervention would be an unfeasible demand for the raters who have a primary responsibility in nursing practice. Deliberation with other members of the nursing team is stimulated, to incorporate as much as possible the experience of the team as a whole. The random selection of raters allocated this intervention level survey to 13 Dutch speaking and 7 French speaking raters. All 20 raters have completed the survey (non response = 0).

7.4 STATISTICAL ANALYSIS

7.4.1 Robust selection of staffing needs

In a first step descriptive measures per nursing intervention were assessed concerning central tendency and distribution. Item N100 of NMDSII, defining blood administration, will be treated as an example.

Table 18: NMDSII item N100: Blood administration, Descriptive measures of estimated nursing time

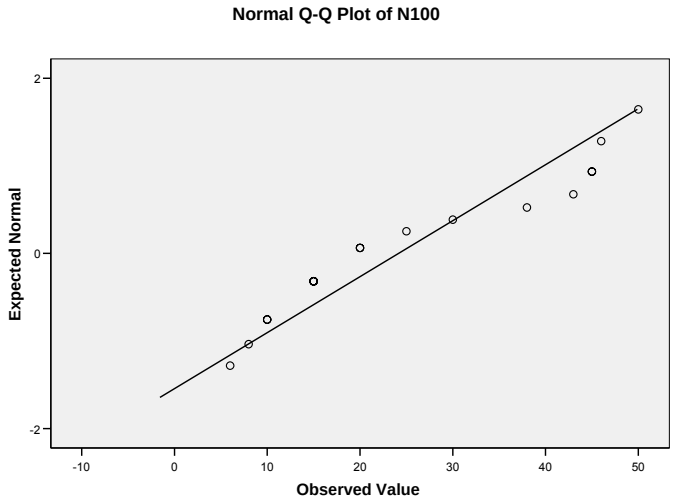
			Statistic	Std. Error
N100	Mean		24,1579	3,58870
	95% Confidence Interval for Mean	Lower Bound	16,6183	
		Upper Bound	31,6975	
	5% Trimmed Mean		23,8977	
	Median		20,0000	

For N100 as an example mean estimated appropriate nursing time equals approximately 24 minutes per administered blood unit. It indicates that if blood administration is indicated, it will take by estimation 24 minutes of nursing time each day to accomplish adequate care. Just as on a patient case level, as discussed in the previous chapter, one notices a distribution slightly skewed to the right and a potential influence of outliers. Both kinds of estimation concern a rating in minutes. And both ratings are biased similarly by some raters who rate much higher than the majority of raters within a normal distribution. For example N100 this is confirmed by the following figures:

Table 19: Normality measures for example NMDSII intervention N100

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
N100	0,195	19	0,057	0,893	19	0,036

Figure 17: Normality figure for example NMDSII intervention N100



As can be seen on the graph above, the skewed trend is not always clear cut caused by random noise retrieved in all data sources. But taking all interventions into account, a right skew ness is apparent.

The same robust selection strategy as on a patient case level is followed. If the Shapiro – Wilk test indicates a significant deviation from normality, Huber's mean is used as an alternative to the mean. For each intervention of NMDSII this correction is applied.

7.4.2 Calibration of relative points and application on a patient case level

Next to analysis on a nursing intervention level, it would be also interesting to aggregate the nursing intervention time needs on a patient case level. 112 patient cases describing nursing care during 24 hours are constructed, as described in the previous chapter. NMDSI and NMDSII were also collected for each of these cases. The combination of NMDSII items in each patient case makes it possible to sum the separate nursing intervention time needs as a result of the survey, to incorporate the whole patient case content using this second alternative method. This makes the findings of both methods, patient case based and intervention level based, directly comparable.

To facilitate interpretation the rating per intervention is recoded to a relative weight, expressed in relative points, before programming. This makes the weights mutually comparable by relating them to the same standard. A denominator of five minutes was chosen. This is comparable to other relative weighting systems and makes a weight to minutes back calculation easier. A relative instead of absolute system is also preferred because else it is tempting to use the number of minutes in daily practice in absolute terms. This should be avoided because relative ratios are the main concern in the current study. An absolute needs estimation falls without the study scope.

Nursing intervention time needs were aggregated on a patient case level by programming the summation of case specific NMDSII weights using SAS version 9.1. Items that are frequency based in NMDSII are taken into account by multiplying their weight with their frequency before summation.

The same aggregation logic was followed using patient case NMDSI data. This enables a comparison with other workload estimation systems based on NMDSI.

7.5 RESULTS

7.5.1 Differential weights of nursing time needs on a nursing intervention level

Nurse time needs will be presented following the structure of NMDSII, based on the Nursing Intervention Classification (NIC). It consists of six main Domains, which are subdivided in Classes. Items of these Classes are grouped in a clinically logical way. Some item groups such as B200 and B400 represent forms of nursing care which exclude each other out within one registration. For example B240 and B250, a urinary stoma and a urinary catheter, can't both be registered on the same patient day. If both forms are present on the same day, e.g. due to a care transition, only the highest code is registered in NMDSII. Secondly, many items show different options in specific care modalities. Here also only one modality can be registered per patient day. And thirdly, items such as B300 are frequency based, as already mentioned before. In interpretation of Nurse time weights of these items it is important to know that these weights first have to be multiplied with item frequency of presence to make them comparable with the others.

Domain I: Care for elimentary fysiological functions					
Class A Support of activities and physical movement					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
A100	Structured physical exercises		60,6	44,9	12
Class B Care for elimination					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
B100_1	Elimination child care	toilet trained child day and night time	50,7	15,9	10
B100_2		toilet trained child night time	40,5	18,8	8
B100_3		non toilet trained child day and night time	30,2	13,9	6
B210	Urinary elimination follow-up		13,5	8,3	2^{H}
B220	Support of urinary continent patient		38,7	30	6^{H}
B230	Care for the urinary incontinent patient		48,7	25,3	10
B240	Care for urinary stoma		32,7	21,9	7
B250	Care for urinary catheter		26,4	17,2	5
B300	Bladder catheterization	x frequency	19,9	9,4	4
B410	Faecal elimination follow-up		13,9	13,3	2^{H}
B420	Support of faecal continent patient		27,8	20,4	5^{H}
B430	Care for the faecal incontinent patient		33,8	20,6	7
B440	Care for faecal stoma or pouch		24,6	11,4	5
B500	Constipation prevention or treament		22,5	14,1	4^{H}
B600	Elimination care education		29,0	15,8	6

H = Huberts mean as estimator, else mean as estimator

Structured physical exercises is attributed a relatively high weight. In B100 the ordinal scale of weights is a bit strange. One would expect an opposite ranking, except if a toilet trained child indeed requires more nursing time than a not toilet trained child. The ranking is clearer in B200 and B400 according to patient dependency.

Class C Care for patient mobility					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
C110	24 h bedridden patient installation		44,8	33,1	7[H]
C120_1	non 24 h bedridden patient installation	supervision	38,5	31,6	6[H]
C120_2		partial assistance	29,7	16,1	6
C120_3		complete assistance	40,5	22,4	8[H]
C200_1	Support of intraward patient mobility	supervision	24,7	16,1	5[H]
C200_2		complete assistance	27,1	17,8	5
C300	Extraward patient transport	x frequency	14,1	5,4	3
C400	Care for traction		30,2	18,7	6
Class D Care for feeding					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
D110_1	Support of feeding (patient room)	supervision	17,3	9,1	3[H]
D110_2		partial assistance	31,5	14,1	6
D110_3		complete assistance	66,4	27,8	13
D110_4		specific feeding needs	67,4	25,3	13
D120_1	Support of feeding (dining room)	supervision	14,5	6,2	3
D120_2		partial assistance	25,0	9,5	5
D120_3		complete assistance	45,1	18,4	9
D120_4		specific feeding needs	76,5	19,2	15
D130	24 h sober patient care		8,3	5,5	2
D200	Care for child bottle and breast feeding		79,3	33,1	16
D300_1	Administration of gastro enteral tube feeding	gastric tube	37,9	19,1	8
D300_2		stoma	39,2	19,9	8
D400	Administration of Total Parenteral Nutrition		21,4	11,8	4

H = Huberts mean as estimator, else mean as estimator

According to the raters there is little difference in nursing time needs within certain items. This is the case for item C120, C200 and D300. Another specific care modality doesn't always imply another level of nurse time needs. Sometimes only content changes.

Class E Comfort support					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
E100	Symptoms management pain		24,3	15,1	5
E200	Symptoms management nausea and emesis		19,8	10,6	4
E300	Symptoms management tiredness		17,4	11,2	3
E400	Symptoms management sedation		24,7	14,2	5
Class F Personal care support					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
F110_1	Hygienic care at lavatory, bed or incubator	supervision	13,9	8,3	3
F110_2		partial assistance	20,3	10,9	4
F110_3		complete assistance	35,8	20,9	6^{H}
F110_4		permanent presence and guidance	24,4	14	5
F120_1	Hygienic care in bath or shower	supervision	15,9	9,9	3^{H}
F120_2		partial assistance	21,8	11,1	4
F120_3		complete assistance	28,8	11,6	6
F120_4		permanent presence and guidance	25,5	9,6	6^{H}
F200	Hygienic care education and training		25,8	12,7	5
F300	Support of day clothing		20,1	14,2	3^{H}
F400	Support of self image		19,4	12	4
F500	Special mouth care		29,8	18,7	6

H = Huberts mean as estimator, else mean as estimator

Class E and F show that the range of attributed nurse time weight is sometimes limited. Three and six are the outer limits here in weighting.

Domain II: Care for complex fysiological functions					
Class G Care for base acid and elektrolyt balance					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
G100_1	Regulation of hydration and nutritional balance	hydration and nutrition 1/day	20,3	12,1	4
G100_2		in out 1/day	18,3	10,8	4
G100_3		in out 2-6/day	27,0	15,7	5
G100_4		in out 7-12/day	37,6	23,9	8
G100_5		in out > 12/day	65,8	42,4	13
G100_6		in out electronic	21,1	7,9	4
G200	Care for evacuating gatric tube		24,4	17,2	5
G300_1	Regulation of glycemic balance	without education	31,7	20,9	6[H]
G300_2		with education	47,3	26,4	9[H]
G400	Regulation of blood balance		27,6	7,8	6
G500_1	Dialysis regulation	peritoneal dialysis	104,8	45,3	21
G500_2		discontinual hemodialysis	95,5	25,5	19
G500_3		continual hemodialysis	73,8	8,5	15
Class H Care for drug use					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
H100	Administration of SC, ID or IM medication	x frequency	11,1	7,7	2[H]
H200	Number of different administered IV drugs	x frequency	11,3	7,2	2[H]
H300	Most frequently administered IV drug	x frequency	7,9	4,1	2
H400	Administration of aerosol, puff or oxygen tent medication	x frequency	11,9	10,1	2[H]
H500	Administration of vaginal medication	x frequency	15,4	8,7	3

H = Huberts mean as estimator, else mean as estimator

Item G100_6 is a good example of the positive influence of the use of technology on nurse time spending. The electronic format permits an almost continual monitoring of an important parameter, while greatly reducing nurse time needs compared to other modalities. In a context of multitasking this however also implies less time to do other nursing intervention simultaneously (e.g. direct monitoring, psycho social support), since direct patient contact diminishes in duration. Notice also the high weighting of dialysis regulation.

Class I Neurological care					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
I100	Neurological function follow-up using GCS		11,4	8,8	1[H]
I200_1	Pressure monitoring of intracranial fluid	without drainage	25,8	5,6	5
I200_2		with drainage	30,4	7,2	6
Class K Care for breathing					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
K100	Aspiration of airways		42,2	30,4	7[H]
K200_1	Supportive means of breathing function	mask, gogles, nasal tube, oxygen tent	12,7	8,6	3
K200_2		endotracheal tube, larynx mask	35,8	13,4	7
K200_3		trachea canulae	60,1	29,8	12
K300_1	Artificial ventilation	regular	115,4	64,1	23
K300_2		special type	104,0	32,4	21

H = Huberts mean as estimator, else mean as estimator

Item K300, artificial ventilation, is attributed a relatively high point. This seems realistic given the high nursing time requirements to care for patients with artificial ventilation. In contrast the weight of item I100, neurological function follow-up, is almost negligible.

Class L Skin and wound care					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
L100	Supervision of wound dressing, materials and near skin		11,7	9,4	2^{H}
L200	Care for sutures and inward materials points	x frequency	17,8	10,4	3^{H}
L300	Simple care for open wound	x frequency	17,6	10,1	4
L400	Complex care for open wound	x frequency	35,5	16,3	7
L500	Care for dermatological leasions	x frequency	19,8	12,2	4
Class M Regulation of temperature					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
M100	Thermal regulation follow-up		54,6	30,3	11
Class N Care for tissue circulation					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
N100	Administration of blood and blood components	x frequency	24,2	15,6	4^{H}
N200	Artificial entry point supervision and/or care	x frequency	11,9	8,3	2^{H}
N300	Venous blood sampling	x frequency	11,3	4,9	2
N400	Arterial blood sampling	x frequency	10,5	5,2	2^{H}
N500	Capillar blood sampling		7,2	5	1^{H}
N600	Cardio circulation support by electrical aids		44,3	12,7	9
N700_1	Cardio circulation support by mechanical aids	internal assistance device	88,6	11,7	18
N700_2		external assistance device	93,1	5,2	19
N700_3		ECMO	58,4	35,2	12

H = Huberts mean as estimator, else mean as estimator

N700 is also a high nurse time weight item. Wound care (L items) and care for tissue circulation (N items) are often frequency based. The frequency of specific care delivery will determine the relative weighting.

Domain III: Behavioral care					
Class O Behavioral therapy					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
O100_1	Activity support	group	60,8	10,5	12
O100_2		individual	26,2	10,9	5
O200	Behavioral dysfunctioning care		64,5	13,2	13
Class P Cognitive therapy					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
P100_1	Care for patients with reduced cognitive abilities	occasional	53,6	43	11
P100_2		standard plan	54,0	35,6	11
Class Q Communication support					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
Q100	Support of communication problems		43,5	31,6	7^{H}
Class R Problem handling support					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
R110	Basic emotional support		30,8	19,8	6
R120	Specific emotional support		53,7	34,1	11
R130	Emotional crisis support		48,0	35,1	10
Class S Patient education					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
S100_1	Specific education	occasional	30,4	22,9	5^{H}
S100_2		standard plan	31,4	18,5	6
S200_1	Pre investigation or surgical procedure education	occasional	31,4	25,5	5^{H}
S200_2		standard plan	29,1	21,3	6

H = Huberts mean as estimator, else mean as estimator

The items concerning behavioural care receive a relatively high weighting. Emotional support is clearly also considered an important aspect of nursing care with its own nurse time needs.

Domain IV: Safety care					
Class V Risk management					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
V100	Pressure ulcer prevention by means of dynamic materials		23,6	16,9	5
V200	Pressure ulcer prevention by repositioning	x frequency	14,2	10,2	2^{H}
V300	Continual monitoring of vital parameters		56,8	32,2	11
V400	Discontinual monitoring of vital parameters		18,3	12,5	4
V500	Tissue or excremental sampling	x frequency	11,7	8,2	2
V600_1	Isolation care	minimal 2 elements of {apron, gloves, mask, garbage handling}	59,8	45	11
V600_2		minimal 3 elements and seperate patient room	54,8	38,6	9
V700	Protective measures with des orientation		44,3	34,4	7

H = Huberts mean as estimator, else mean as estimator

At first continual monitoring of vital parameters (V300) and isolation care (V600) seem to stick out in nurse time weighting. However, items such as repositioning (V200) and tissue sampling (V500) are frequency based.

Domain V: Family care					
Class W Birth care					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
W100	Relaxation care in preparation of child birth		238,4	103,6	48
W200	Ante partum care: monitoring uterine activity	x frequency	67,0	46,3	13
W300	Child birth delivery		83,0	13,5	17
W400	Post partum follow-up	x frequency	15,3	5,8	3
W500	Kangooroo care		40,1	17,9	8
Class X Family care					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
X100	Rooming in of family or significant others		29,6	19,7	4

H = Huberts mean as estimator, else mean as estimator

Item W100, ante partum care, has a weight of 48, which is extreme compared to the other weights. This seems to point to a biased estimation. Maybe the raters considered all time spent ante partum by the patient in the hospital as a form of ante partum care. Or maybe the questioning in the survey was unclear.

Domain VI: Healthcare management					
Class Y Care counseling					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
Y100	Cultural brokerage		21,0	12,2	4
Y200	Anamnesis at intake		19,6	11,5	4
Class Z Management of care provisions and information					
item	**Care description**	**Specific care modality**	**Nurse time mean**	**Nurse time SD**	**Nurse time weight**
Z100	Functional, mental, psychosocial assessment		31,0	21	6
Z200	Physician support in direct medical care	x frequency	41,7	30,1	7
Z300	Multidisciplinary conference		32,6	19,2	6
Z400	Contact with other institutions		12,5	7,4	2

H = Huberts mean as estimator, else mean as estimator

Healthcare management receives relatively low weighting in nursing care. Back office aspects such as patient record administration needs and continuous informal contacts between caregivers are however not included.

7.5.2 Comparison with workload estimation systems based on NMDSI

Two other validated nursing workload weighting systems were compared to the constructed relative point system: the use of 'Points_closon' and 'Points_gent'. The three weighting systems were calculated based on NMDSI and NMDSII of the 112 constructed patient cases, as described in the previous chapter.

As you can see in the correlation table below, the three systems are highly correlated to each other. These correlations, together with all previous positive consistency checks, prove that 'Relative_points' on a 24 hours level as a basis for further model construction is a well founded option.

Table 20: Alternative relative points correlation matrix

		Relative_points	Points_closon	Points_gent
Relative_points	Pearson Correlation	1	,928(**)	,945(**)
	Sig. (2-tailed)		0,000	0,000
	N	112	112	112
Points_closon	Pearson Correlation	,928(**)	1	,992(**)
	Sig. (2-tailed)	0,000		0,000
	N	112	112	112
Points_gent	Pearson Correlation	,945(**)	,992(**)	1
	Sig. (2-tailed)	0,000	0,000	
	N	112	112	112

**. Correlation is significant at the 0.01 level (2-tailed).

7.5.3 Time validation

It is also very interesting to compare the ratings of patient cases as a whole, which was the outcome described in the previous chapter, with the ratings of nursing interventions as described in this chapter. To make both comparable nursing intervention ratings were aggregated based on NMDSII of the patient cases.

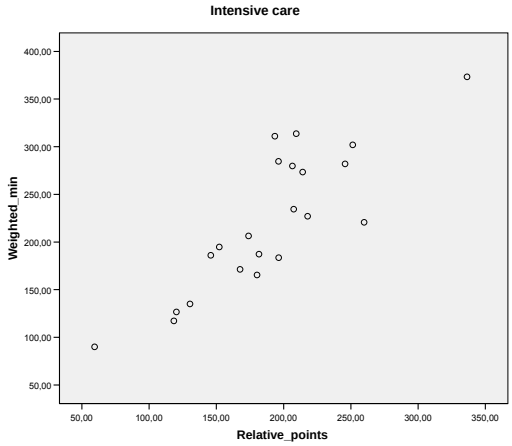
Table 21: Patient case – nursing intervention level correlation matrix

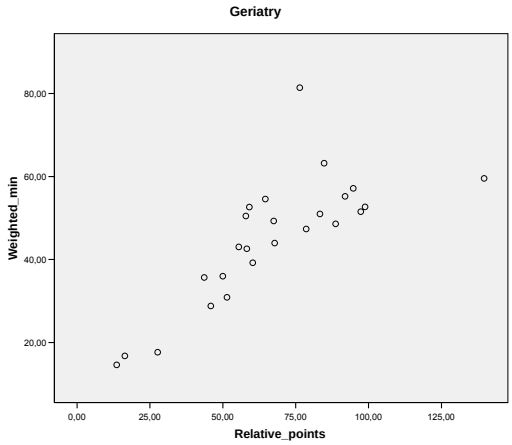
		Patient case rating	Relative points
Patient case rating	Pearson Correlation	1	,902(**)
	Sig. (2-tailed)		0,000
	N	112	112
Relative points	Pearson Correlation	,902(**)	1
	Sig. (2-tailed)	0,000	
	N	112	112

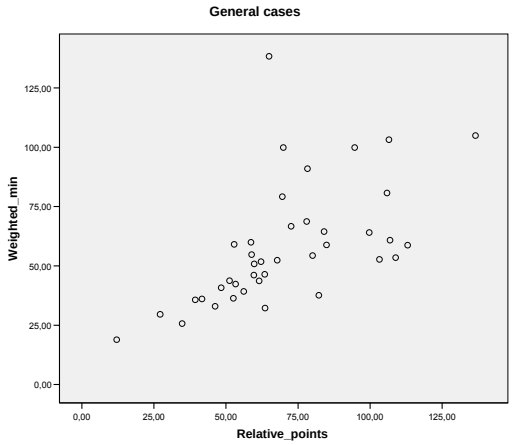
**. Correlation is significant at the 0.01 level (2-tailed).

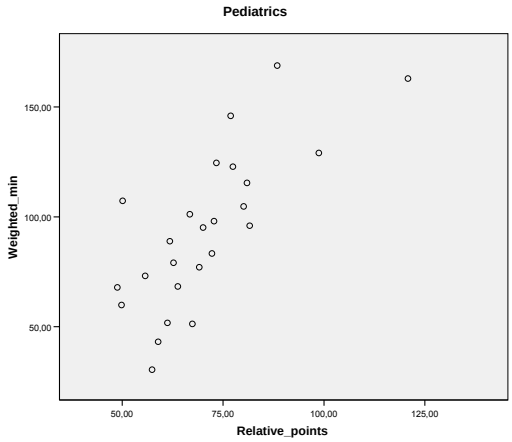
There is a very high correlation between the ratings of patient cases as a whole on a 24 hours basis and the summation of ratings of separate nursing interventions based on NMDSII. Keep in mind that both ratings were performed independently by other charge nurses not knowing anything of the alternative rating. Both methods were also aimed at a very different level of nursing care, giving clinical nursing information in a very different format. So both estimate approaches are highly consistent with each other, with NMDSI based systems and with accurate external patient classification systems (see previous chapter).

The high correlation between the relative points and the patient case nurse time measure is also confirmed visually, as presented per specialty ICU, Geriatrics, General care and Paediatrics:



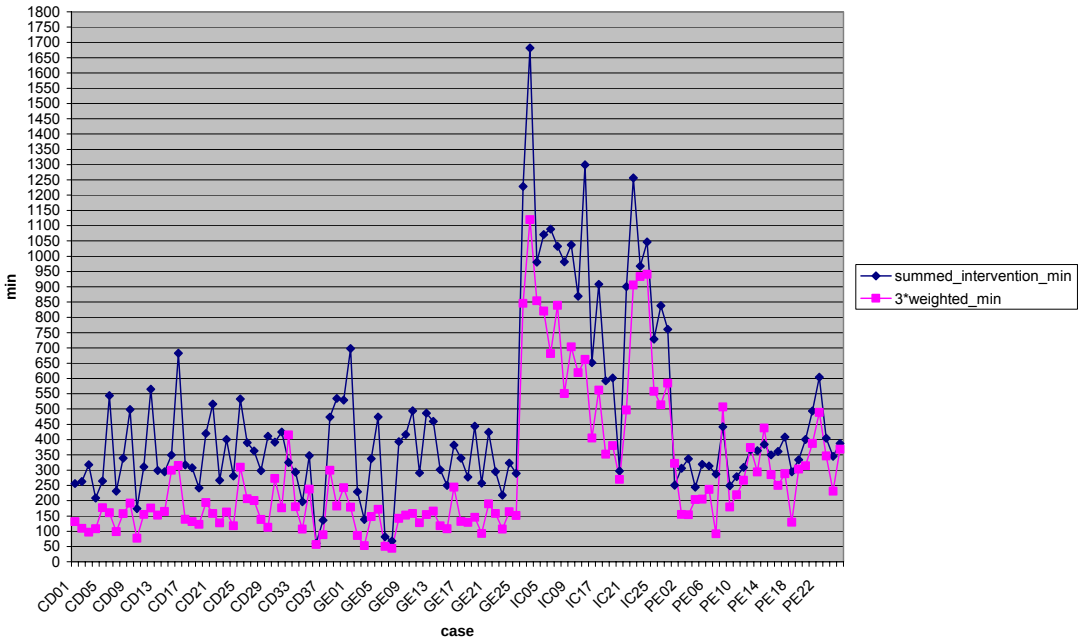






The figure below presents the estimation for each patient case on a 24 hours basis. Cases are ordened on the horizontal axis by specialty. The first group represent General care (CD), secondly Geriatrics (GE), then ICU (IC) and finally Paediatrics (PE). The blue line presents relative points as summation of intervention nurse time weights. The pink line represents the direct patient case rating on a 24 hours basis.

Figure 18: Comparison of estimated nursing time needs by case



Intensive care cases are consistently rated higher in nursing time needs than general, paediatric and geriatric cases. The differentiation between the other specialties is less clear. The starting hypothesis is confirmed: summing intervention estimates per case (mean 465, SD 291) leads to an consistent surplus compared to estimating the time for a patient case as a whole (mean 285, SD 229). So there is a mean difference of 180 min (SD=130) in estimation of nurse time needs between both approaches concerning 24 hours of patient stay. This difference is smallest in paediatric cases and highest in intensive care cases. Only in 5 patient cases summed intervention ratings fall below total case ratings. In 96% of patient cases hypothesis is confirmed. Trend consistency between both estimates is also confirmed visually.

7.6 DISCUSSION AND CONCLUSIONS

Next to a patient case method also an activity based method is feasible to assess nurse time needs to deliver quality nursing care. A relative point system weighting nursing interventions was constructed. A robust selection of nurse time estimation was required to account for skew ness and influential observations.

NMDSII is a very useful tool to assess nurse time needs on a nursing intervention level. It is coherent and structured in its registration of nursing care. Many patient factors such as level of dependency are already incorporated in NMDSII. It has the features of classic patient classification systems, but it is a national and hospital wide system. It is endorsed by sector participation and statistical analysis. And it is a dataset that can be linked to other relevant datasets such as HDDS.

The weighting of nurse time needs per nursing intervention seem intuitively and clinically logical. Some aberrations such as the weighting of pre partum relaxation (W100) should be investigated further.

There is a reassuringly high correlation with NMDSI based weighting systems Closon and Ghent. A further validation by comparison with real time nursing intervention time measurement can be useful. Although a direct comparison with the PRN workload estimation method was not possible for reasons stated in the previous chapter, the proposed use of NMDSII relative weights corresponds to the PRN methodology. Both methods attribute relative points to required nursing interventions. Each point corresponds to five minutes of nursing time. NMDSII remains a bit more flexible in practical execution terms of intervention definitions. The standard NIC framework is however applied, using a minimal set of quality ensuring requirements to register an intervention. The planning in terms of length of duration of intervention application is left to the discretion of the healthcare team involved, as opposed to PRN. Rather a maximal approach was applied: Providing sufficient time to enable quality of care was the main target. The definition of 'quality of care' was considered context specific and therefore concretized based on each participant's professional expertise. NMDSII is based on 79 nursing interventions, where as PRN 80® uses 214 'actions'. PRN integrates direct and indirect care (i.e. communication, mobility, administration). Communication and mobility aspects are partially included into NMDSII. Administration and back office in general is considered too organization specific to assess nationwide in a comparable way. Continuous education of caregivers and quality improvement initiatives are also not included in NMDSII. The latter two aspects of care are however very important in providing good patient care. A NMDSII based financing system should always be complemented with additional resource allocation concerning back office, staff education and quality improvement.

The hypothesis that intervention times aggregation would lead to a surplus when compared to ratings on a patient case level was confirmed. But the correlation between both methods is strikingly high and unexpected.

8 STAFFING IMPLICATIONS OF THE USE OF EVIDENCE BASED NURSING

8.1 INTRODUCTION

Mott et al (2005)[125] found that 43% of nurse respondents (n = 99) were unable to identify a source of information and resources about Evidence Based Practice. The use of EBN isn't wide spread in nursing practice. Unawareness, the lack of an evaluative culture and the lack of organizational support and management commitment to the development of Evidence Based Nursing are barriers to implementation and further regular use [126].

The relation between Evidence Based Practice (EBP) and staffing issues is currently limited in research to effect studies evaluating the relevance of staffing quantity and qualifications for clinical outcome in terms of mortality and morbidity. Staffing is recognized as an important contributing factor in delivering the ultimate healthcare goal: providing high quality care.

But it's unclear how high the number of staffing should be. For now historical ratios drive these decisions. In addition, the much larger field of clinical process knowledge (as opposed to outcome), treated in guidelines, algorithms, etc., isn't considered in staffing research. So, a pertinent question surfaces: can EBP also be used as a tool to determine the appropriate number of staff in delivering nursing care? Considering that recommended clinical processes are proven to lead to beneficiary outcomes, this seems a valid reasoning. This study part tries to yield first premature insights into the feasibility of adaptation of staffing quantity based on EB clinical recommendations. Patient care needs are used as the primary driver of staffing needs. The rating procedure, as described in chapter 4, will be used to evaluate a potential difference in staffing needs.

8.2 RESEARCH QUESTION

This study part is aimed at the following research questions:

Is there a difference in estimated nursing staff needs, and their resource cost, between the actual care versus evidence based care approach?

If yes, what is the nature of this difference?

8.3 METHODOLOGY

8.3.1 Case selection for EBN modification

The methods of general case selection and construction are described extensively in Chapter 4. Therefore we briefly summarize this approach in resume: In order to assess staffing and resources needs as linked to nursing care, a large group of raters evaluated the actual staffing needs in real patient cases. 112 clinical cases were constructed for the rating procedure. A described patient case consist of a written story of a patients day, NMDSI scores, NMDSII scores and a specialized scoring like San Joaquin (index C,D, H*), TISS-28 (index I), Narvel (index E) and AGGIR (index G).

Based on the frequency of the EBN documented NMDSII-items as present in the 112 patient cases, five were selected for EBN modification. In table 22 frequencies of NMDS items per selected case are presented.

Table 22: The patient cases with the highest frequency of previously generated E.B. knowledge

Casus	pressure ulcer prevention: dynamic systems	pressure ulcer prevention: changing positions	protective measures	pain management	isolation	special mouth care	urinary incontinence	urinary catheter	artificial ventilation	Number of EB inter-ventions
16		X	X	X		X		X		5
25		X		X	X	X	X			5
26	X	X				X	X	X		5
38	X	X	X	X		X		X		6
40	X	X	X	X	X	X		X		7

As you can see, most of these cases combine more than one of the nine examined E.B. nursing interventions.

8.3.2 Modification of patient cases

The modification of the cases is rolled out in phases. In a first phase each of five reviewers, specialized in one to three EB nursing intervention(s), screened all five cases on indications and contraindications for this (those) specific nursing intervention(s). In a second phase every reviewer consulted all highlighted relevant evidence independently, as applicable on the selected cases. The responsible reviewer of phase I proposed adaptations of his interventions for every case. After a discussion modifications were accepted or rejected based on hard EBN rules or consensus of the team members in case of ambiguity. Some methodological problems had to be tackled in this process:

- Indications and contra indications: Does the collected case information, based on real patient records, offer sufficient ground to warrant or prohibit the use of a nursing intervention or a modification in its execution? Contextual data is given in the patient case descriptions. However, a brief case description can never hold the same level of information as presented in real life care and complete patient records which are available to caregivers. So, in the real life cases additional relevant information could have been present that guided the originally described care in the cases. The retrospective modifications can't take this absent information into account. This potential for information lacking can influence the justification for modifications in multiple ways. Some modifications in reality could be unjustified; other necessary modifications could be overlooked. Same hazard holds on the level of modifications in execution of nursing interventions. However, the goal of assessing the difference between an EB versus non EB staffing approach can be attained independently of this research limitation. In the comparison of ratings of EB versus non EB cases it is assumed that the implemented modifications are justified. The described and rewritten patient cases are considered as the basis for comparison, not the original real patient situations. This doesn't influence the objectivity of findings concerning the presence or absence and degree of difference in estimated staffing needs between EB versus non EB. Like in other parts of the study, relative estimations offer sufficient grounding instead of absolute real needs. The latter falls outside the scope of the study.

- The time element: A case consists of one day of nursing care during the patient stay. Evidence based recommendations concern the whole stay or don't specify a unique timing within the patient stay. In practice not all elements are done on every day of stay. A realistic set of evidence based interventions for one day must be identified, taking into account that other care aspects are already done before or will be done afterwards. The considered clinical timeframe (e.g. admission, postoperative care, day of hospital discharge) is of a significant influence. The same reasoning as mentioned above holds to counter this research limitation. Artificially or not, the comparison in rated nurse time needs remains unbiased.

- The individualistic versus standardized approach: The set of inserted intervention evidence consists of elements that are standardized over individual patients (e.g. the prohibition of use of protective measures for patients with higher risks on fall accidents as a sole indication). Other elements have to be individualized. In the

modified cases the assumption was made that there was no ground present for an exceptional approach. This again is a step away from clinical reality, because exceptional need to modify evidence based guidance is not uncommon.

- Interdependence between conflicting guideline and/or systematic review evidence wasn't present in the implemented comparison. Some authors have reported difficulties in combining different sets of guideline evidence in real patient cases due to co morbidity effects on the integrated treatment needs. For example, a blind patient with diabetes shouldn't receive eye drops, although this is prescribed in regular diabetic treatment guidelines. The same holds for many medication interactions which are unaccounted for in the implementation of each guideline separately in a real patient case. This potential problem wasn't encountered in the present study, but could potentially bias future replications. The fact that medication isn't the main topic in evidence based nursing and nursing interventions diminishes the probability of such bias when compared to evidence based medicine and its application.

In a third phase, EBN interventions' protocols were described in detail according to EB – guidelines. These protocols describe in detail every modification that has been made in each of the five selected cases, and its evidence based rationale. This process is for one nursing intervention illustrated in Appendix 10.

In Appendix 11 an example of a case modification is given. As you can see some parts are highlighted in color. These paragraphs represent modification concerning mouth care, isolation care, symptoms management pain, urinary incontinence care and pressure ulcer prevention.

8.3.3 Delphi survey of EBN cases

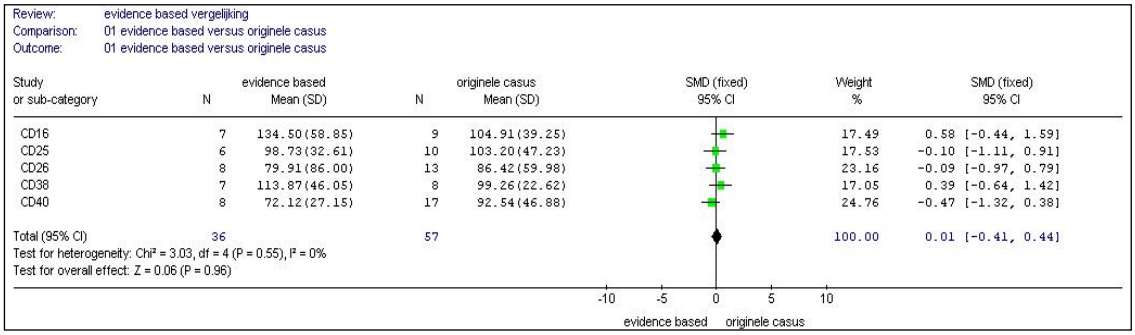
After adapting the cases a Delphi survey was performed. In total 10 raters received all 5 EB adapted cases via email. The methodology of this Delphi study was exactly the same as the Delphi survey on the cases without modification. Raters had to fill in their ratings via an interactive website and received feedback of their rating and the ratings of other raters later on by email. After two rounds data collection ended. The response-rate in total was 72% on case-level. On rater-level the response-rate was less: 62%.

8.3.4 Comparison of EBN adapted cases versus Non EBN adapted cases.

EBN rating data is compared with the original results of the first Delphi survey using forest plot methodology in the EB statistical program Review Manager®. This method takes mean as well as variance into account in computing a potential difference.

8.4 RESULTS

Figure 19: Forest plot of EB versus non EB case ratings



In figure 19 results are reported per selected patient case. No difference is found between the EBN transformed cases and non-EBN transformed cases in estimated nursing time needs. There is also no clear direction of effect. When the five cases are analysed together, no significant difference was found.

8.5 DISCUSSION AND CONCLUSIONS

Based on the implemented comparison with its current limitations no difference could be found with regard to the use of evidence based methods or not to define the nursing time needs in specific clinical situations.

During the preparation of the evidence based cases it was remarkable that most modifications implied an addition of extra components or action points concerning the selected interventions. Only in a few exceptions an actual activity had to be deleted, because it was clearly unwarranted. Therefore, as hypothesis one would expect that evidence based estimated staffing needs would exceed experience based estimated staffing needs. The raters judged it otherwise. There is no clear cut tendency, nor is it significant.

Some considerations have to be made about these results. Firstly only five cases are examined. It is possible that a difference can be found when a broader perspective of nursing is exposed to a rating panel. Secondly the number of respondents per case is rather low. The modification of the cases was only based on a few nursing interventions. When other or more interventions are manipulated, the effect could be more substantial.

Another important consideration concerns the limits in what can be changed in a written patient case based on evidence. As already mentioned, one is limited to the background information presented in the original case. This case holds a set of clear indications and contra indications. These facilitate modification. In contrast, the evidence guidance also refers to many other additional clinically relevant indications and contra indications, each of which should be examined in surplus of the original information. This knowledge gap can only be resolved based on direct patient case experience. This is not possible in a retrospective design.

As mentioned, replication of this part of the study is recommended using more cases, interventions and raters using a prospective study design. For now, there is no evidence which supports a significant effect of the use of guidelines, systematic reviews and other evidence based guidance to determine justified staffing needs instead of estimated actual needs.

Considering the relevant amount of resources in time and costs to implement a modification of such a limited scope, one can argue that a widespread use of this approach is questionable. By means of research it can be a way to stimulate healthcare professionals to apply evidence based methods in daily practice. The EB versus non EB cases are very clear to confront treatment difference between own and recommended activities. A peer review application is one possibility.

9 DEVELOPING A NURSE TIME NEEDS MODEL AS A BASIS FOR FINANCING

9.1 INTRODUCTION

Hughes (1999)[116] considers the statistical methods as a separate approach in treating nursing intensity. Statistical methods base their predictions on previous information from the clinical area itself. Information is collected on significant patient characteristics such as the age, length of stay, etc. Regression analysis is carried out to identify appropriate nursing time for different permutations of patient characteristics [116]. However, as extensively illustrated in chapter 1, variability in nursing care is poorly explained by the currently predominantly medical data such as registered in HDDS. The opportunity presents itself to investigate this within the Belgian NMDSII context. The chapter 4 study part provided an indicator of nurse time needs on a patient case level. Chapter 5 determined nursing time weights per NMDSII nursing intervention in an independent manner. A very high correlation was found between both measures. Backed by high concurrent validity findings, it is possible to use the nursing intervention time weights on the whole of nationally collected NMDSII data. Treating the programming of weighting to a 24 hours level as the dependent variable (Y), the contribution of independent variables (X's) to the dependent variable can be assessed. HDDS and/or NMDSII can be implemented as independent variables into the model. Based on the level of explained variance in the prediction of nurse time needs, a model can be constructed as a basis for financing hospital nursing care.

9.2 RESEARCH QUESTION

This study part is aimed at the following research questions:

- Is it possible to construct a model predicting nurse time needs based on HDDS and/or NMDSII?
- How much nurse time needs variance is explained?
- What is the potential for further use in the finance scheme of hospital nursing care?

9.3 METHODOLOGY

9.3.1 Preparation of a HDDS – NMDSII national minimal dataset

NMDS-II information was collected during the pilot study of actualisation NMDS-II and consists of 117.395 inpatient days from 66 Belgian hospitals. Each hospital participated in the test with a minimum of 2 and a maximum of 5 nursing wards. A balanced sample was obtained for the following medical specialties: geriatrics (index G), pediatrics (index E), intensive care (index I), chronic illness (index SP), maternal services (M), and general internal medicine (index D, H*) and general surgical procedures (index C, H*)

From these hospitals the MKG data for the corresponding hospitalization units and registration period (year 2003 semester 2 and year 2004 semester 1) was collected. From 59 hospitals the quality of the data was sufficient to create a dataset of DRG's using the apr-drg grouper from 3M.

The NMDS-II dataset also contained registrations of inpatient days from newborns. Because newborns admitted with their mother on index « M », have no separate stay in MKG, these inpatient days where excluded. NMDS-II and DRG information was then matched using CTI / CIV code of the hospital and the identification code of the hospital stay. This resulted in a dataset « DRG-MVG » of 66.827 inpatient days from 18.148 hospital stays, with DRG and NMDS-II information.

Still some supplementary MKG information was needed. Therefore a dataset was constructed out of « DRG-MVG » with one record per stay ("gekoppelde_verblijven"). Following information was matched from MKG_stayhosp: patient number, MKGtypestay, type_admission, Length Of Stay, MKG_facdays, age_indicator and year of admission. Based on patient number and CTI / CIV code, year of birth was looked up from MKG_pathhospi to calculate the age of the patient at the year of admission.

For every surgical stay the first day of surgery was determined. Therefore two datasets were constructed based on the 3M definitions manual of APR-DRG version 15.0. One dataset with all operating-room procedures and a second one with most frequently used non-operating-room procedures that lead to grouping a hospital stay into a surgical APR-DRG. These two procedure-datasets were matched with MKG_procicd9 and the first day of every stay was determined as the day of surgery. These days of surgery where then matched with the surgical stays in "gekoppelde_verblijven".

To group the APR-DRG's in Fetter groups of "similar nursing intensity", a mapping had to be done between HCFA-DRG's – AP-DRG's and APR-DRG's. A frequency table was used to match APR-DRG's to their most frequently corresponding AP-DRG's and their Fetter group. The subdivision of DRG within Fetter groups can be consulted in Appendix 12.

All these supplementary MKG information was matched with the NMDS-II information in the dataset « DRG-MVG ». Based on date of admission and date of registration (NMDS-II) the relative day of registration was calculated. The relative day of admission is « 1 ». For every surgical stay the relative day of registration minus the day of surgery lead to a new value where the relative day on the day of surgery is « zero ».

From these dataset the stays with MKGtypestay "H" where selected to create database "MKGtype H werktabel". This dataset consists of 11.670 stays and has 60.019 records.

Finally index information from MKG_stayndx had to be matched with every day of registration. Therefore a dataset ("patbeweeg") was constructed where every day of every stay (MKGtypestay = "H") is represented by one record. This record contains: CTI / CIV code, identification number of the stay, index and relative day of stay. To construct this dataset LOS per index was used. When the number of factdays was 1 higher than sum of LOS, one day was added at the end of the stay.

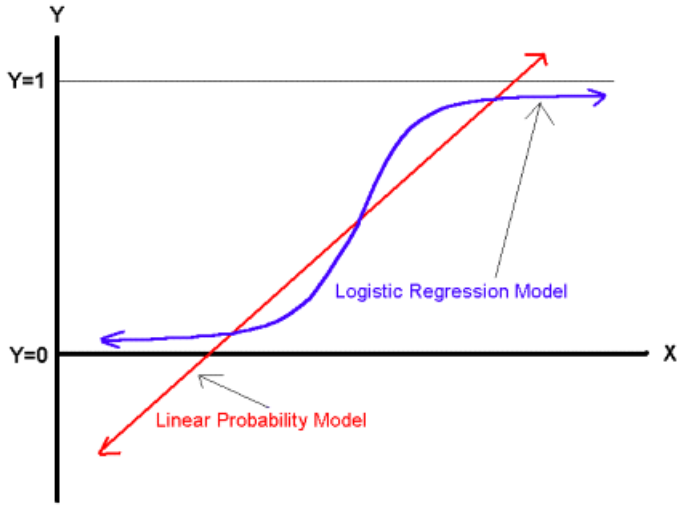
In a last step the Relative points as indicator of nurse time needs were calculated for each patient record, based on the programming of nursing intervention weights on NMDSII on a 24 hours level.

The dataset was randomly divided in two datasets A and B with respectively 30.009 records (9.926 stays) and 30.010 records (9.981 stays), one for modeling purposes and one for testing purposes.

9.3.2 Modelling approach

In a next step it is important to predict 'relative_points' as representation of relative staffing needs by selecting the most significant, but feasible subset of predictors in the reference dataset. Generalized Linear Modelling (GLM) was tried as a first approach to build an accurate model. A GLM stepwise reduction regression was implemented, including all two way interaction effects. This GLM modelling led to an explained variance of only 26,4%. Multinomial logistic regression was considered a more appropriate approach.

Logistic regression is based on the assumption that a logistic relationship (i.e. a sigmoidal dependency) exists between the probability of group membership and one or more predictor variables. If there are two groups, binary logistic regression is used, whereas if there are three or more groups, a choice has to be made between nominal and ordinal logistic regression. Nominal logistic regression is used when there is no natural ordering to the groups, whereas ordinal logistic regression is used when there is a rank ordering. Logistic regression can be used to predict a dependent variable on the basis of independents and to determine the percent of variance in the dependent variable explained by the independents; to rank the relative importance of independents; to assess interaction effects; and to understand the impact of covariate control variables.



Logistic regression applies maximum likelihood estimation after transforming the dependent into a logit variable (the natural log of the odds of the dependent occurring or not). In this way, logistic regression estimates the probability of a certain event occurring. Note that logistic regression calculates changes in the log odds of the dependent, not changes in the dependent itself as ordinary least squares (OLS) regression does.

Ordinary linear regression determines the relationship between a continuous outcome variable and the predictor variables. Logistic regression determines the relationship between the probability of the outcome occurring and the predictor variables. Logistic regression has many analogies to OLS regression: logit coefficients correspond to b coefficients in the logistic regression equation, the standardized logit coefficients correspond to beta weights, and a pseudo R^2 statistic is available to summarize the strength of the relationship. Unlike OLS regression, however, logistic regression does not assume linearity of relationship between the independent variables and the dependent, does not require normally distributed variables, does not assume homoscedasticity, and in general has less stringent requirements. The success of the logistic regression can be assessed by looking at the classification table, showing correct and incorrect classifications of the dichotomous, ordinal, or polytomous dependent.

The multinomial logistic regression model is a generalization of logistic regression to outcomes with more than two levels. The model is also known as polytomous or polychotomous logistic regression in the health sciences and as the discrete choice model in econometrics [127]. It is a multiequation model.

9.3.2.1 Definition of the model

Logistic regression analysis extends the techniques of multiple regression analysis to research situations in which the outcome variable is categorical.

Suppose we were considering the simple linear regression model where the response or dependent variable is binary. The following model illustrates this situation.

$$Y_i = \beta_0 + \beta_1 X + \varepsilon_i; \quad Y = 0, 1; \quad E(\varepsilon_i) = 0$$

Hence,

$$E(Y_i) = \beta_0 + \beta_1 X = \pi_i.$$

We see then that when the response variable is binary, the expected value of the response variable is a probability. Specifically, it is the probability that Y=1.

Unfortunately, there are several difficulties associated with this model.

Non-normal Error Terms: Since $\beta_0 + \beta_1 X$ is a fixed quantity, Yi is a random variable strictly because □i is (or visa versa).

Since Yi takes on only two values (0 and 1), □i also takes on only two values.

$$\varepsilon_i = 1 - \beta_0 - \beta_1 X \text{ when } Y_i = 1.$$
$$\varepsilon_i = -\beta_0 - \beta_1 X \text{ when } Y_i = 0.$$

Clearly, there can be no assumption of normality distributed error terms in this instance.

Nonconstant Error Variance: Since Yi and □I differ by only a constant, they have the same variance.

$$Var(Y_i) = Var(\varepsilon_i) = \pi_i(1-\pi_i)$$
$$= (\beta_0 + \beta_1 X_i)(1 - \beta_0 - \beta_1 X_i).$$

The important thing to note here is that the error variance is not constant; it depends on the level or value of Xi.

Constraints on the Response Function: Since the response function represents a probability for a binary dependent variable, the mean responses should be constrained by,

$$0 \le E(Y_i) = \pi_i \le 1.$$

We see that the linear model is applicable when the outcome variable, Y, is continuous, but is not appropriate for situations in which Y is categorical. For example, if Y takes on the value 1 for "success" and 0 for "failure," the multiple regression equation would not result in predicted values restricted to exactly 1 or 0. In fact, these predicted values would be spread out over an interval that contains uninterpretable values such as .5 or .33 and could even include negative values and/or values greater than 1. The model for logistic regression analysis, described below, is a more realistic representation of the situation when an outcome variable is categorical.

Solution:

However, many response functions will not naturally possess this property. That is, values of Xi within the reasonable range of interest may produce predicted responses outside the interval [0, 1].

The problems of non-normality and non-constant variance of the error terms might be handled by weighted least-squares or some parameter estimation techniques which are not sensitive to normality. However, the problem with the constraints cannot be easily addressed in the context of linear regression. It is necessary to consider a nonlinear model whose response function has the property of asymptotically approaching 0 (on the left) and 1 (on the right) over the range(s) of the independent variable(s).

One widely used response function is the logistic response function. For simple regression, these have the form,

$$E(Y_i) = \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)} = \left[1 + \exp(-\beta_0 - \beta_1 X)\right]^{-1}.$$

An important property of the logistic function is that it can be linearized. For example,

$$\pi' = \ln\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 X.$$

This transformation is sometimes called the logit transformation; π' is called the logit mean response and the ratio $\frac{\pi}{1-\pi}$, is called the odds.

Note that π' is defined over (- ∞ , + ∞)

When several predictors are used to model the response, it is convenient to express the model in the familiar form:

$$E(Y_i) = \left[1 + \exp\left(-\boldsymbol{\beta}^T \mathbf{X}\right)\right]^{-1}, \text{ and } \pi' = \ln\left(\frac{\pi}{1-\pi}\right) = \boldsymbol{\beta}^T \mathbf{X},$$

where, $\boldsymbol{\beta}^T \mathbf{X} = \beta_0 + \beta_0 X_1 + \beta_0 X_2 + \cdots + \beta_0 X_{p-1}.$

Likelihood Function

Since the responses, Yi, are independent, the joint pdf is

$$g(Y_1, Y_2, \ldots, Y_n) = \prod_{i=1}^{n} \pi_i^{Y_i} (1 - \pi_i)^{1-Y_i}.$$

Hence,

$$\ln g(Y_1, Y_2, \ldots, Y_n) = \sum_{i=1}^{n} Y_i \ln\left(\frac{\pi_i}{1-\pi_i}\right) + \sum_{i=1}^{n} \ln(1 - \pi_i).$$

Now, since

$$\pi_i = E(Y_i) = \left[1 + \exp\left(-\boldsymbol{\beta}^T \mathbf{X}\right)\right]^{-1},$$

It follows that

$$1 - \pi_i = \left[1 + \exp\left(\boldsymbol{\beta}^T \mathbf{X}\right)\right]^{-1}.$$

Substituting

$$\ln\left(\frac{\pi_i}{1-\pi_i}\right) = \beta^T \mathbf{X},$$

we write the log-likelihood as

$$\ln L(\beta) = \sum_{i=1}^{n} \left(Y_i \, \beta^T \mathbf{X} - \ln\left[1 + \exp\left(\beta^T \mathbf{X}\right)\right] \right)$$

The Multinomial Logistic Regression Model

The multinomial logistic regression model is a generalization of logistic regression to outcomes with more than two levels.

Suppose the multinomial outcome variable Y takes values in the set {1, . . . , g}. The multinomial logistic regression model assumes that the probability for observation i to have outcome s depends on i's covariates x_{i1}, . . . , x_{ip} as

$$P(Y_i = s) = \frac{e^{\eta_{is}}}{\sum_{t=1}^{g} e^{\eta_{it}}},$$

where

$$P(Y_i = s) = \frac{e^{\eta_{is}}}{\sum_{t=1}^{g} e^{\eta_{it}}},$$

is a linear predictor. In this formulation of the model we have a regression coefficient β_{ks} for each combination of covariate k and outcome category s, and a separate linear predictor ηis for each outcome category [127].

A response variable with k categories will generate k-1 equations. Each of these k-1 equations is a binary logistic regression comparing a group with the reference group. Multinomial logistic regression simultaneously estimates the k-1 logits. For example, if we have a dependant variable with three levels, the probabilites for each of the levels could be obtained as follows:

P(y=1) = exp(β1*x)/(exp(β1*x)+exp(β2*x)+exp(β3*x))

P(y=2) = exp(β2*x)/(exp(β1*x)+exp(β2*x)+exp(β3*x))

P(y=3) = exp(β3*x)/(exp(β1*x)+exp(β2*x)+exp(β3*x))

This system of equations is unidentified, that is, there is more than one solution to the coefficients that lead to the same probabilities. To make the system identifiable, one of the coefficients is set to 0. It doesn't matter which one since they each yield the same probabilities. We will set the probability for β1 to 0, yielding:

P(y=1) = 1/(1+exp(β2*x)+exp(β3*x))

P(y=2) = exp(β2*x)/(1+exp(β2*x)+exp(β3*x))

$$P(y=3) = \exp(\beta 3^*x)/(1+\exp(\beta 2^*x)+\exp(\beta 3^*x))$$

This, in turn, leads to the following probabilities relative to the reference group, in this case, group 1.

$$P(y=2)/P(y=1) = \exp(\beta 2^*x)$$

$$P(y=3)/P(y=1) = \exp(\beta 3^*x)$$

Thus, the two coefficients, β2 and β3 represent the log odds of being in the target groups relative to the reference group.

In multinomial logistic regression the relative risk can be defined as,

rr1 = P(y=1)/P(base category)
rr2 = P(y=2)/P(base category)
...

Thus, the relative risk ratio for multinomial logit would be

$$RRR = \frac{P(y=1|x+1)/P(y=\text{base category}|x+1)}{P(y=1|x)/P(y=\text{base category}|x)}$$

Hosmer and Lemeshow (2000)[127] suggested looking at the multinomial model as if it were a set of independent ordinary logistic models of each outcome against the reference outcome, and testing the fit of each of these separately. Lesaffre and Albert (1989)[128] give diagnostics for detecting influential, leverage and outlying samples in multinomial logistic regression.

The properties of the resulting test are verified using simulated data and illustrated on a liver enzyme data set in Albert and Harris (1987) [129].

9.3.2.2 Significance Tests

The likelihood ratio test is based on -2LL (deviance). The likelihood ratio test is a test of the significance of the difference between the likelihood ratio (-2LL) for the researcher's model minus the likelihood ratio for a reduced model. It is an alternative to the Wald statistic, and is also called the log-likelihood test.

Test of the overall model :

The model fitting information gives the results of iterations after minimization of criteria like Akaike Information Criterion (AIC)[a], Bayesian Information Criterion (BIC) and -2 times the log of the likelihood function (-2LL)[b]. Thus the likelihood ratio test of a model tests the

---

[a] The AIC is an approximately unbiased estimator for a risk function based on the Kullback–Leibler information.

[b] A "likelihood" is the probability that the observed values of the dependent may be predicted from the observed values of the independents. Like any probability, the likelihood varies from 0 to 1. The log likelihood (LL) is its log and varies from 0 to minus infinity (it is negative because the log of any number less than 1 is negative). LL is calculated through iteration, using maximum likelihood estimation (MLE). Log likelihood is the basis for tests of a logistic model.
The likelihood ratio is a function of log likelihood. Because -2LL has approximately a chi-square distribution, -2LL can be used for assessing the significance of logistic regression, analogous to the use of the sum of squared errors in OLS regression. The -2LL statistic is the likelihood ratio and reflects the significance of the unexplained variance in the dependent.

difference between -2LL for the full model and -2LL for initial chi-square in the null model. This is called the the model chi-square test. The null model, also called the initial model, is logit(p) = the constant. That is, initial chi-square is -2LL for the model which accepts the null hypothesis that all the b coefficients are 0. This implies that that none of the independents are linearly related to the log odds of the dependent. Model chi-square thus tests the null hypothesis that all population logistic regression coefficients except the constant are zero. It is an overall model test which does not assure that every independent is significant.

Degrees of freedom in this test equal the number of terms in the model minus 1 (for the constant). This is the same as the difference in the number of terms between the two models, since the null model has only one term. Model chi-square measures the improvement in fit that the explanatory variables make compared to the null model. Model chi-square is a likelihood ratio test which reflects the difference between error not knowing the independents (initial chi-square) and error when the independents are included in the model (deviance). When probability (model chi-square) <= .05, we reject the null hypothesis that knowing the independents makes no difference in predicting the dependent in logistic regression.

Test of individual model parameters :

The likelihood ratio test assesses the overall logistic model but does not tell us if particular independents are more important than others. This can be done, however, by comparing the difference in -2LL for the overall model with a nested model which drops one of the independents. We can use the likelihood ratio test to drop one variable from the model to create a nested reduced model. In this situation, the likelihood ratio test tests if the logistic regression coefficient for the dropped variable can be treated as 0, thereby justifying dropping the variable from the model. A nonsignificant likelihood ratio test indicates no difference between the full and the reduced models, hence justifying dropping the given variable so as to have a more parsimonious model that works just as well.

9.3.2.3 Pseudo R-Square

The « pseudo » term comes from the fact that we cannot made direct analog to OLS regression's R-square. R^2 measure seeks to make a statement about the "percent of variance explained," but the variance of a categorical dependent variable depends on the frequency distribution of that variable. This means that R-squared measures for logistic regressions with differing marginal distributions of their respective dependent variables cannot be compared directly with R^2 from OLS regression is also problematic. Nonetheless, a number of logistic R-squared measures have been proposed.

Cox and Snell's R-Square is an attempt to imitate the interpretation of multiple R-Square based on the likelihood, but its maximum can be (and usually is) less than 1.0, making it difficult to interpret.

Nagelkerke's R-Square is a further modification of the Cox and Snell coefficient to assure that it can vary from 0 to 1. That is, Nagelkerke's R^2 divides Cox and Snell's R^2 by its maximum in order to achieve a measure that ranges from 0 to 1. Therefore Nagelkerke's R-Square will normally be higher than the Cox and Snell measure. It is the most-reported of the R-squared estimates. [130]

In appendix 13, multinomial logistic regression a table of "Parameter Estimates" with k-1 tiered sections is presented, where k= the number of categories of the dependent. The last (kth) category is omitted as a reference category. Each tier will have a row for the intercept, each continuous variable, and each dummy value of each categorical variable. One of the columns will be the odds ratios, labelled "Exp(b)." The larger odds ratios within a tier indicate which variables have the most effect for that tier's category of the dependent variable.

9.3.2.4 Assumptions

Logistic regression does not assume a linear relationship between the dependents and the independents. The dependent variable need not be normally distributed. The dependent

variable need not be homoscedastic for each level of the independents; that is, there is no homogeneity of variance assumption. Normally distributed error terms are not assumed. However, logistic coefficients will be difficult to interpret if not coded meaningful, error terms are assumed to be independent (independent sampling). Violations of this assumption will occur, for instance, in repeated measures designs. To the extent that one independent is a linear function of another independent, the problem of multicollinearity will occur in logistic regression, as it does in OLS regression. As the independents increase in correlation with each other, the standard errors of the logit (effect) coefficients will become inflated. Multicollinearity does not change the estimates of the coefficients, only their reliability. High standard errors flag possible multicollinearity.

9.3.3 Modelling inputs

The goal of the statistical model is to study how the variables resulting from the Hospital Discharge Data Set (HDDS) can predict the time of nursing intervention. This variable was calculated on the basis of the Minimum Nursing Data Set II (MNDS II) by applying times dedicated for each item NMDS and estimated by the experts. To be applicated in multinomial logistic regression, we will thus be brought to discretize the independent variable with representative categories; in the same way it will be necessary to categorize the independent variables when they are quantitative.

During the registration of MNDS II, continuous measurements were realized on consecutive days. Thus, we can find the same patient several times in the data base and this more especially for the long stays. Over the 59.904 recorded days, we raise 11.649 (19.4%) distinct stays, and consequently 48.255 (80.6%) repeated patients. The following table is obtained:

Table 23: Sequential count of matching cases

Count	Frequency	Percent	Cumulative Percent
1	11649	19.45	19.45
2	10090	16.84	36.29
3	7824	13.06	49.35
4	6199	10.35	59.70
5	4997	8.34	68.04
6	3598	6.01	74.05
7	3028	5.05	79.10
8	2664	4.45	83.55
9	2226	3.72	87.26
10	1961	3.27	90.54
11	1302	2.17	92.71
12	1125	1.88	94.59
13	980	1.64	96.23
14	888	1.48	97.71
15	773	1.29	99.00
16	94	0.16	99.16
17	84	0.14	99.30
18	74	0.12	99.42
19	69	0.12	99.53
20	60	0.10	99.63
21	29	0.05	99.68
22	29	0.05	99.73
23	29	0.05	99.78
24	23	0.04	99.82
25	23	0.04	99.86
26	19	0.03	99.89
27	19	0.03	99.92
28	18	0.03	99.95
29	17	0.03	99.98
30	12	0.02	100
31	1	0.00	100
Total	59904	100	

A model based on repeated measures is not appropriate because of the effect of the patients whose stay is longer and of this fact will have a too important weight. Moreover, one of the criteria of application of the model is to avoid the repeated data to respect the assumption of independent sampling.

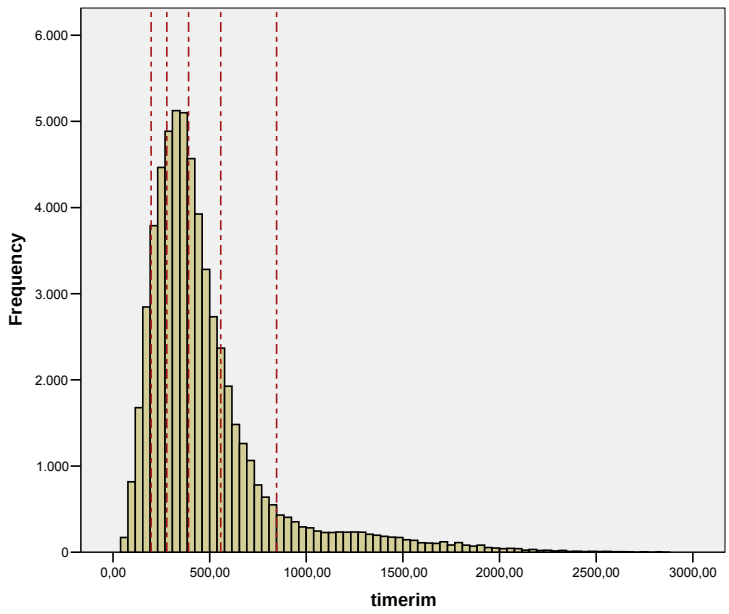
To build an aggregate from hospitalization days asks to choose among the repeated data of the same patient that which represents the case as well as possible. In the model below, we chose the maximum time care cost. On the basis of this subset of data, we built a predictive model by means of the multinomial logistic regression.

9.3.3.1 Dependent variable (TimeRimCod)

Previously, we can carry out the categorization of the variable for its use in the multinomial logistic regression. A first categorization based on the percentiles will be applied as we can see in the next table. This categorization takes account of the Gaussian character of the distribution. One can however test another method of discretization while cutting the statistical series in bands of same size. This method postulates that we are in the presence of a multinomial uniform distribution, which is not really the case. We have tested the two methods but the first one is the best.

Table 24: Categorization of TimeRimCod

TimeRimCod		N	Percent	Mean	Std. Deviation	Minimum	Maximum
1	< P10	5 939	10%	151.7	33.6	44.3	197.1
2	P10 - P25	8 909	15%	239.8	23.1	197.1	278.2
3	P25 - P50	14 850	25%	334.8	32.1	278.2	391.5
4	P50 - P75	14 847	25%	464.1	46.9	391.5	556.8
5	P75 - P90	8 909	15%	664.2	78.5	556.8	845.8
6	> P90	5 939	10%	1 310.6	374.2	845.9	2 863.8
Total		59 393	100%	481.6	338.8	44.3	2 863.8



9.3.3.2 Independent variables

The predictors are the six Fetter's DRGs (Vplk_cluster), severity of illness (SOI), Risk of mortality (ROM), Bed index (Indx), type of admission (TypAdm), sex, six categories of lenght of stay (LosCod), six categories of stay ratio [c] (StayRatioCod), six categories of age (AgeCod), and if the drg is surgical (P) or medical (M) (APR_DRG_PM).

9.3.3.3 Case Processing Summary

Variables	Modalities	N	Marginal Percentage
TimeRimCod	1	613	5.50
	2	1341	12.04
	3	2604	23.37
	4	3091	27.74
	5	2044	18.35
	6	1448	13.00
Vplk_cluster	1	1463	13.13
	2	2301	20.65
	3	2247	20.17
	4	2237	20.08
	5	2861	25.68
	6	32	0.29
SOI	1	2835	25.45
	2	3730	33.48
	3	2957	26.54
	4	1619	14.53
ROM	1	5266	47.27
	2	2397	21.52
	3	2241	20.11
	4	1237	11.10
Indx	CDH*	5536	49.69
	E	2656	23.84
	G	1381	12.40
	SI	759	6.81
	Sp	809	7.26
TypAdm	Urg	5999	53.85
	Plan	5142	46.15
sex	M	5742	51.54

[c] We can express the moment of the stay like the relationship between the day of registration and the total duration of the stay.

Variables	Modalities	N	Marginal Percentage
	V	5399	48.46
LosCod	1	2984	26.78
	2	2164	19.42
	3	2811	25.23
	4	1845	16.56
	5	898	8.06
	6	439	3.94
StayRatioCod	1	2400	21.54
	2	433	3.89
	3	3090	27.74
	4	2609	23.42
	5	1501	13.47
	6	1108	9.95
AgeCod	1	1765	15.84
	2	2017	18.10
	3	2991	26.85
	4	2405	21.59
	5	1221	10.96
	6	742	6.66
APR_DRG_PM	M	7964	71.48
	P	3177	28.52
Valid		11141	100.00
Missing		508	
Total		11649	

9.4 RESULTS

The results show that fitting criteria are significant. All the dependent variables are significant except sex and the type of admission. The Nagelkerke's pseudo-R square gives an percentage of explication of 0.41. This result is confirmed by the classification matrix which gives a global correct classification percentage of 39.3%. We can see that it is the low categories of times that are less predicted.

At this time of the study, we cannot extrapolate estimated times on hospitals. The data collected at the time of the study of actualization of the MNDS II do not constitute a random sample. The case mix of the study is very different from the national case mix. Because of the selection of care programs, some apr-drg shows a poor sample size, and conversely, others are overestimated.

9.4.1 Model Fitting Information

	Model Fitting Criteria			Likelihood Ratio Tests		
Model	**AIC**	**BIC**	**-2 Log Likelihood**	**Chi-Square**	**df**	**Sig.**
Intercept Only	32146.3	32182.9	32136.3			
Final	26893.0	28137.2	26553.0	5583.2	165	0.00000

9.4.2 Pseudo R-Square

Cox and Snell	0.39416
Nagelkerke	0.40817

9.4.3 Likelihood Ratio Tests

Effect	**Model Fitting Criteria**			**Likelihood Ratio Tests**		
	AIC of Reduced Model	**BIC of Reduced Model**	**-2 Log Likelihood of Reduced Model**	**Chi-Square**	**df**	**Sig.**
Intercept	26893.0	28137.2	26553.0	0	0	.
Vplk_cluster	26951.4	28012.5	26661.4	108.4	25	2.3736E-12
SOI	27009.9	28144.3	26699.9	146.9	15	9.9021E-24
ROM	26967.1	28101.4	26657.1	104.1	15	2.2019E-15
Indx	28195.9	29293.6	27895.9	1342.9	20	1.955E-272
TypAdm	26892.3	28099.8	26562.3	9.2	5	0.10059132
sex	26893.5	28101.0	26563.5	10.4	5	0.06372401
LosCod	27001.0	28062.2	26711.0	158.0	25	2.7926E-21
AgeCod	27308.7	28369.9	27018.7	465.7	25	9.6129E-83
StayRatioCod	27233.6	28294.8	26943.6	390.6	25	2.6536E-67
APR_DRG_PM	27021.5	28229.0	26691.5	138.4	5	3.8311E-28

9.4.4 Classification table

	Predicted Response Category							
Observed	**1**	**2**	**3**	**4**	**5**	**6**	**Total**	**Percent Correct**
1	13	17	393	151	12	27	613	**2.12**
2	8	26	723	484	54	46	1 341	**1.94**
3	12	20	1 217	1 059	191	105	2 604	**46.74**
4	5	9	844	1 668	412	153	3 091	**53.96**
5	4	5	274	916	583	262	2 044	**28.52**
6	2	1	118	227	229	871	1 448	**60.15**
Total	**44**	**78**	**3 569**	**4 505**	**1 481**	**1 464**	**11 141**	**39.30**
Overall Percentage	**0.39**	**0.70**	**32.03**	**40.44**	**13.29**	**13.14**		

9.4.5 Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	5101.1	25	0
Likelihood Ratio	4205.5	25	0
Linear-by-Linear Association	2950.6	1	0
N of Valid Cases	11141		

9.4.6 Statistics of association

			Value	Asymp. Std. Error	Approx. T	Approx. Sig.
Directional Measures	Somers' d	Symmetric	0.43771	0.00703	59.93	0.00000
		TimeRimCod Dependent	0.46948	0.00748	59.93	0.00000
		Predicted Response Dependent	0.40996	0.00670	59.93	0.00000
Symmetric Measures		Kendall's tau-b	0.43872	0.00705	59.93	0.00000
		Gamma	0.57006	0.00853	59.93	0.00000
		Spearman Correlation	0.50663	0.00792	62.02	0.00000
Measure of Agreement		Kappa	0.21206	0.00583	44.23	0.00000

9.5 DISCUSSION AND CONCLUSIONS

The consistency of the constructed basic nurse time measure has been proven based on multiple comparable trend estimators. Therefore this measure fits its purpose. Building a nursing financing model based on staffing needs, is possible within a broad set of scenario's to tackle specific issues. The treatment of data on a shift level and its further aggregation, the choice of an appropriate main staffing needs measure, correction for skew ness and outlier influence; … For each of these issues different options are tested and appropriate solutions were found.

This resulted in a grounded system of relative weighting of staffing needs, which is totally based on NMDSII, collected on a national level. It is an appropriate staffing and nursing resource estimator on a patient day of stay level. It's also very transparent in resource allocation.

However, one of the main goals of the study was also to investigate the relationship with HDDS and APR – DRG. One of the weak points in current financing is the calculation of relative staffing/resource needs solely based on NMDSII is its risk of data 'creep', i.e. optimization and/or manipulation. Therefore, a model based on HDDS data, as proposed, could reduce the effect of NMDSII data manipulation.

This has been conducted on a patient day of stay level. All fitting criteria, except sex and type of admission were significant, i.e. Fetter groups, SOI, ROM, bedindex, length of stay, age, stayratio and surgical versus medical distinction. The model shows that HDDS explains only 39 to 41% in relative staffing/resource needs.

However, the minimal data sample used was not representative for DRG distribution on a national level. Therefore this study part has confirmed the feasibility of constructing a NMDSII based nursing finance model with a whole or partial link to HDDS data. Further analysis is premature, since a representative sample of minimal data should first be obtained.

10 GENERAL DISCUSSION AND CONCLUSIONS

This hospital nursing financing study was a feasibility study. A wide range of alternatives were evaluated, among which some seem appropriate for further investigation.

The Belgian hospital nursing financing system is regarded as one of the systems which are fairer to nursing care because it takes nursing data into account for calculation of the hospital reimbursement. Not many countries do this. In most countries, the average nursing cost per day is taken into room and board costs. It means that nursing costs are directly correlated to a length of stay independent from the real nursing needs of patients. There is high evidence that nursing care needs are varying from patient to patient and from day to day. There is evidence that the costs directly attributable to the last day of a hospital stay are an economically insignificant component of total costs. Reducing LOS by as much as 1 full day reduces the total cost of care on average by 3% or less [131].

The primary disadvantage of this approach is that there is a cost compression issue, e.g. hospitals that have low nursing intensity patients do better within this reimbursement framework, and hospitals that have high nursing intensity patients tend not to do as well [55]. Compression refers to the tendency of all DRG weights to be too close to the average when measurement is not sensitive enough to accurately capture actual cost differences across DRGs. It may be expected that a chief contribution to compression results from incomplete measurement of patients' use of the largest component of inpatient care: nursing services [132].

There are a limited number of countries such as Canada, Australia, New Zealand, and Switzerland in which nursing workload data are taken into account for calculating cost weights per DRG. This leads to a more detailed process of data capturing, but much closer to the real differences in costs. The main issue in most countries is not the question if the reimbursement system should be adjusted for nursing data, but rather the availability of the data. Most countries, such as the USA, see the added value, but are resistant to the fact that the data are systematically available in a comparable way.

Belgium is seen as one of the leading countries in which a nursing minimum data set is available and the hospital reimbursement system is adjusted for nursing care. The issue is in the way the adjustment is done. It has been seen as not transparent and not linked with DRGs, too much weighted on actual instead of appropriate nurse staffing data, and too sensitive to data creep and manipulation because the system is built on registration of presence or frequency of performed nursing interventions. The main goal of this study was to investigate alternatives to meet these requirements.

A first major conclusion of the study is that it is possible to weight nursing care based on an appropriate staffing level instead of actual staffing levels.

In the study 112 real clinical patient cases were written, judged on staffing needs by clinical nurses and head nurses. The questions were quite simple: if you had to care for these patients, how much time would it take? How much of these patients are you able to care for? If you wouldn't have any limitations on resources, would that make a difference?

These cases were randomly distributed among nurses so that every nurse had to rate on average about 10 cases and every case was evaluated on average by 8 nurses. These nurses didn't know the patient in question, nor were they working in the hospital in which the case was written. The nurses who were rating these cases came from 69 different hospitals. At the same moment of writing the case, the nursing minimum dataset (I and II) and some relevant patient classification systems were scored.

The result is firstly that there is a high internal consistency among raters between the different questions which shows a high reliability. Secondly, there is a high correlation among these ratings and the scores from patient classification systems such as TISS, AGGIR, and San Joaquin which shows external consistency. The conclusion is that these ratings are useful to calibrate nursing time and costs. At the same time there is a high variability among raters in assessing the time needed to care for these patients. There was a small, but significant difference in case ratings between the Dutch and French speaking region.

The differences among raters are not unexpected. These differences are linked to differences in physical environment, working conditions, organization, skill mix, staffing perceptions etc. More than 100 of these determinants have already been described by Young in 1981. Robust statistical methods have been used for estimating the most appropriate nursing time needed for rendering quality healthcare in each of these clinical cases.

Independent of the rating of the cases, each of the 79 nursing interventions of the nursing minimum dataset II was rated by 20 randomly chosen raters, evaluating the time needed to perform each of these interventions separately. Again a high variability of time ratings was seen between raters. And in an identical way, robust estimators were used to calculate a time average for each nursing intervention.

Because of the availability of the nursing minimum dataset for each case, all weights per scored intervention were added. The sum score of interventions per case correlated for more than 91% with the independent nurse time ratings per case as a whole. The high correlation between these scores indicates a high concurrent validity of the obtained scores. Although there is a wide difference among raters, the estimated average time (as well for cases as for individual interventions) seems to be highly valid. It offers sufficient grounding for using these time weights in further applications. There is a need for external validation of these time weights. The cases, which exist in Dutch and French, could be easily rated in Dutch and French speaking countries such as The Netherlands, France, Switzerland, and Luxembourg to evaluate if the differences between Dutch and French speaking raters are significant in a broader European perspective. Some nursing interventions have not been scored in the 112 cases, in which they couldn't be evaluated properly. These comments and improvements are minor. There is an outstanding fact that the weighting of the nursing minimum dataset is feasible, reliable and valid.

A second conclusion is that evidence in nursing practice is limited. For the 9 nursing interventions that have been researched, most of the evidence that was found is on level C. There is limited evidence available on level A or B. It means that building the hospital reimbursement system on evidence-based nursing practice instead of actual nursing practice will be difficult due to the scarce evidence available. The question is if it makes a difference on the staffing level required. Five cases were therefore rewritten. For the nursing interventions out of the nine that were researched for evidence, the actual care was replaced by what would be required if all evidence (A to C) was included. Cases were rewritten from 40% to 60%. These 5 cases were given to another 10 random selected raters (different from the ones who had rated the original cases). The ratings were not significantly different. It means that from a nurse staffing viewpoint, there is no real difference in staffing for actual care versus evidence-based nursing care. This is not unexpected. It means that evidence-based care is not always more or less costly. Sometimes some interventions can be avoided. Sometimes some extra interventions will be necessary. Evidence-based care seems to be rather cost neutral. It also means that staffing decisions are probably less precise and not evaluated in some minutes more or less, but rather on caring one more or one less patient.

Interesting is that out of the literature review on pressure ulcers, a decision tree could be developed that can be used as a rule set for querying the hospital and nursing datasets on appropriateness of pressure sore interventions. The main conclusion of a limited search on the available database was that there was more evidence on under care than over care in the Belgian hospitals on pressure sore prevention.

The further development of this rule sets in which nursing and medical data are combined, are an interesting perspective for further research in how more evidence-based care could be included in the hospital reimbursement scheme. The link with pay for performance (P4P) or pay for quality (P4Q) is here most obvious.

A third conclusion is that nursing data can be linked to DRGs. In other countries, a fixed average nursing cost weight per DRG is calculated. The Belgian data are not structured in a way that facilitates calculation of a valid nursing cost weight per DRG. The main issue is the sample design in which only 15% of all inpatient days are collected. It means that 85% of all other days have to be estimated. Using the 15% sample would create an overestimation of longer length-of-stay. In the study, we tried to estimate the required nursing time based on

data that were systematic collected in the hospital discharge dataset (MKG). The variables chosen have been suggested by John Thompson in 1992, being DRG and severity of illness, age of the patient, intensive versus non-intensive care, surgical versus non-surgical care, routine versus emergency care, length-of-stay and the precise day in stay. A multinomial logistic model was used. The estimated time was grouped in 6 groups of nursing time. DRGs were grouped in six nursing relevant classes, given the fact that DRGs can be different from a medical viewpoint e.g. hernia repair and appendectomy but not different from a nursing point of view. The main result is an explained variance of about 40% which is quite encouraging and much higher than found in the literature. The main critique is that the actual model is overestimating low cost care and underestimating high cost care. More crucial is that the model has been built on the available dataset that has been used during the pilot study of developing nursing minimum dataset II. That dataset is not representative for the whole of Belgium from a reimbursement point of view. Only a selection of nursing wards per hospital participated in the study and not all medical care programs where included. The impact of the model on a complete hospital financing couldn't be tested. It would be necessary to test the model on the coupled Belgian hospital discharge and nursing minimum dataset (RCM/MKG – RIM/MVG) for 3 consecutive years. The actual financing scheme needs to be compared with the new developed model.

A major concern is that many nursing data are auto correlated within one stay, within one unit, within one hospital. Taking this into account requires more sophisticated modelling.

A fourth conclusion is that the result of the study, being 6 classes of nursing cost per DRG, is more transparent for users and policy makers than the actual financing methods using zones, ZIP/ZAPs, deciles etc. The statistical background to derive these six classes in the most appropriate way are quite complex, but the result is quite easy to read and understand. Each hospital can compare its own nursing profile per DRG with the national profile. If more rule-sets could be developed and run on the data, evidence-based profiles could be produced that could help hospitals to compare with a more "EBN" benchmark. The actual hospital nursing reimbursement scheme doesn't give any incentive to change practice. The link with DRGs and EBN would help to provide more quality and efficiency incentives.

A fifth conclusion is that linking DRGs and nursing data would help to implement a nursing adjustment to hospital reimbursement on a hospital wide scale. In the actual financing scheme, the adjustment for nursing is limited for surgical, internal medicine, intensive care and pediatric nursing ward. There is no adjustment for geriatric nursing ward although nursing care is one of the major characteristics of patient care on geriatrics. Linking to DRGs makes the reimbursement less dependent on structures and departments and will provide a shift to financing patients and care programmes.

A sixth conclusion is there are many alternatives on how to integrate the nursing component into the hospital reimbursement scheme. A first tentative approach is that nursing data are used for yearly calibration of DRG nursing cost weights. It means that the 15% existing real nursing data are used for estimating the model of 100% of all stays. The main advantage of this approach is that there is no direct impact of scoring on financing so that registration creep for nursing care will be limited. Indeed, over- or underscoring will lead to an adjustment of the model that will apply for all hospitals. At the other hand, it will not hamper registration creep on MKG/RCM which could lead to a higher reimbursement. The main disadvantage is that an average national cost weight per DRG is calculated instead of a nursing cost weight that is hospital specific. It might be that one hospital has a higher cost weight for a given DRG than another hospital because their length of stay is shorter and therefore more compressed and intense or that they tend to have patients with higher nursing needs.

The main concern is that the actual model only explains 40% of the differences in nursing costs and is therefore not sensitive enough to grasp all differences.

A second approach could be that the actual nursing profile per DRG per hospital is taken into account. This approach is more sensitive to creep, but probably more close to differences in nursing practice. The main disadvantage is that we are unsure how the actual 15% sample represents all 100% of nursing care per DRG. Probably, high volume DRGs will be well represented. This is less clear concerning the representation of low volume DRGs and outlier cases.

The final reimbursement scheme could be a mix of modelled and actual nursing care.

The major limitation of this study is that the data on which the model has been built was not representative for the entire hospital healthcare. Further validation of the model is required. Different alternatives and their impact of hospital reimbursement should be investigated on a representative sample of hospital data.

The main value of the study is that the nursing minimum dataset II has been validated for use within the Belgian hospital reimbursement system and the different alternatives of linking DRGs and nursing data have been explored.

11 REFERENCES

1. Sermeus W, Delesie L, Michiels D, Van den Heede K, Van Herck P, Van Landuyt J. Actualisatie van de minimale verpleegkundige gegevens. Eindrapport. Brussel: Ministerie van Sociale Zaken, Volksgezondheid en Leefmilieu; 2006.
2. Delesie L, De Becker P, Sermeus W. Pilootstudie registratie minimale verpleegkundige gegevens. Leuven: Katholieke Universiteit Leuven, Centrum voor Ziekenhuiswetenschap; 1986.
3. Kennedy HP. Enhancing Delphi research: methods and results. J Adv Nursing. 2004; 45(5): 504-511.
4. Welton JM, Zone-Smith L, Fischer MH. Adjustment of inpatient care reimbursement for nursing intensity. Policy Polit Nurs Pract. 2006 Nov; 7(4): 270-80.
5. Sermeus W, Weber P, Chu S, Fischer W, Hunstein D. The DRG Imperative: overview and nursing impact. In Weaver C. et al. (Eds). Nursing Informatics for the 21st Century: An international Look at Practice, Trends and the Future. Chicago: Healthcare Information and Management Systems Society (HIMSS); 2006. 231-245.
6. Fetter RB, Shin Y, Freeman JL et al. Case mix definition by diagnosis-related groups. Medical Care. 1980; 18: iii, 1-iii, 53.
7. Fischer W. Diagnosis Related Groups (DRGs) und Pflege : Grundlagen, Codierungssysteme, Integrationsmöglichkeiten. Bern: H. Huber; 2002.
8. Harrell JS. Predicting nursing care costs with a patient classification system. In: Shaffer FA., Eds. Patients & purse strings: patient classification and cost management. New York: National League for Nursing; 1986. 149-164.
9. Atwood JR, Hinshaw AS, Chance HC. Relationships among nursing care requirements, nursing resources, and charges. In: Shaffer FA (Eds.). Patients & purse strings: patient classification and cost management. New York: National League for Nursing; 1986. 99-120.
10. Green J, Mc Clure M, Wintfeld N, Birdsall C, Rieder KA.. Severity of illness and nursing intensity: going beyond DRGs. In: Scherubel JC & Shaffer FA., eds. Patients and purse strings II. New York: National League for Nursing; 1988. 207-230.
11. Halloran EJ. Nursing workload, medical diagnosis related groups, and nursing diagnoses. Research in Nursing and Health. 1985; 8(4): 421-433.
12. McKibben RC, Brimmer PF, Galiher JM, Hartley SS, Clinton J. Nursing costs and DRG payments. American Journal of Nursing. 1985; 85(12): 1353-1356.
13. Mowry MN, Korpman RA.. Do DRG reimbursement rates reflect nursing costs? Journal of Nursing Administration. 1985; 15(7,8): 29-35.
14. Sovie M, Tarcinale M, Vanputte A, Stunden A.. Amalgam of nursing acuity, DRG's and costs. Nursing Management. 1985; 16(3): 22-42.
15. Wolf GA, Lesic LK.. Determining the cost of nursing care within DRGs. In: Shaffer FA. Eds. Patients & purse strings: patient classification and cost management. New York: National League for Nursing; 1986. 165-180.
16. Bargagliotti LA, Smith H. Patterns of nursing costs with capitated reimbursements. Nurs econ. 1985; 3 (5): 270-275.
17. Dahlen AL, Gregor JR. Nursing costs by DRG with an all-RN staff. In: Shaffer FA., ed. Costing out nursing: pricing our product. New York: National League for nursing; 1985. 123-134.
18. Fosbinder D. Nursing costs/DRG: a patient classification system and comparative study. J Nurs Adm. 1986; 16 (11):18-23.
19. Reschak GLC, Biordi D, Holm K, Santucci N. Accounting for nursing costs by DRG. J Nurs Adm. 1985; 15 (9): 15-20.

20. Riley W, Schaefers V. Costing nursing services. Nurs Manage. 1983; 14(12): 40-43.
21. Walker L, Avant K.. Strategies for theory construction in nursing. Norwalk, CT: Appelton-Centry-Crofts; 1983.
22. Trace LD. The total cost of nursing care for patients with acquired immune deficiency syndrome. In: Scherubel JC, Shaffer FA, eds. Patients and Purse Strings II. New York: National League for Nursing; 1988. 231-248.
23. Welton JM, Halloran EJ. Nursing diagnoses, diagnosis-related group, and hospital outcomes. J Nurs Adm. 2005 Dec; 35(12): 541-9.
24. Thompson JD, Averill RF, Fetter RB. Planning, budgeting, and controlling--one look at the future: case- mix cost accounting. Health Services Research. 1979; 14:111-125.
25. Werley HH, Lang NM, Westlake SK. The nursing minimum data set conference: Executive summary. Journal of Professional Nursing. 1986; 2: 217-222.
26. Werley HH, Devine EC, Zorn CR, Ryan P, Westra BL. The nursing minimum data set: abstraction tool for standardized, comparable, essential data. American Journal of Public Health. 1991; 81: 421-426.
27. Thompson JD, Diers D. Nursing Resources. In Fetter BF, Brand DA & Gamache D. DRGs. Their Design and Development. Ann Arbor: Health Administration Press; 1991. 121-183.
28. Caterinicchio RP, Davies RH. Developing a client focused allocation statistic of inpatient nursing resource use: an alternative to the patient day. Soc Sci Med. 1983; 17(5): 259-272.
29. Diers D, Bozzo J. Nursing resource definition in DRGs. RIMS/Nursing Acuity Project Group. Nursing Economics. 1997; 15: 124-30, 137.
30. Nightingale F. Notes on Nursing. New York: Dover Publications; 1860.
31. Goossen W, Coenen A, Sermeus W, Marin H, Hovenga E, Casey A. The International Nursing Minimum Data Set (i – NMDS). In: Weaver C, Delaney C, Weber P, Carr R, eds. Nursing and Informatics for the 21st century. An international look at practice, trends and the future. Chicago: Healthcare Information and Management Systems Society (HIMSS); 2006. 305-320.
32. Anderson B, Hannah KJ. A Canadian Nursing Minimum Data Set: a major priority. Can J Nurs Admin. 1993; 6: 7-13.
33. Goossen W, Delaney C, Sermeus W et al. Preliminary results of a pilot of the international Nursing Minimum Dataset (i – NMDS). In: Proceedings of Medinfo 11th World Congres on Medical Informatics of the International Medical Informatics Association; 2004. S103.
34. Turtiaenen AM, Kinnunen J, Sermeus W, Nyberg T. The cross-cultural adaptation of the Belgian Nursing Minimum Dataset to Finnish nursing. J Nurs Manag. 2000; 8: 281-290.
35. Berthou A, Junger A. Nursing data. Rapport final: étude des travaux 1998 – 2000. Lausanne: ISE Institut de santé et d'economie. 2000.
36. Paira A, Pereira F. Nursing Minimum Dataset in Portugal. ESEnf.S.Joao, College of Nursing. Available from: www.fnaee.pt/index.php?set=links-esenfs
37. Elo S. The Public Health Nursing Minimum Dataset: Development, application and use. Vard I Norden. 1995; 15(3): 9-16.
38. Scott A et al. Irish NMDS as part of research project, 2003. Available from: www.dcu.ie/nursing/decision_making.shtml
39. International Council of Nurses. ICN Workforce Forum Report, 2003 29th – 30th; September, Oslo, Norway; 2003.
40. Germain B. Le résumé de soins infirmiers, utopie ou nécessité? Objectif Soins. June 1993; 14: 50-8.
41. Picone D, Ferguson L, Hathaway V. NSW Nursing costing study. Sydney Metropolitan Teaching Hospitals Nursing Consortium; 1993.
42. Sermeus W. De Belgische ziekenhuisfinanciering ontcijferd. Leuven: ACCO; 2006.

43. European Healthcare Management Association (EHMA), Healthbasket: Health benefits and service costs in Europe, A European Research Project, April 2004 – March 2007. Available from: http://www.ehma.org/projects
44. Schreyogg J, Stargardt T, Tiemann O, Busse R. Methods to determine reimbursement rates for diagnosis related groups (DRG): a comparison of nine European countries. Health Care Manag Sci. 2006 Aug; 9(3): 215-23.
45. Busse R, Schreyogg J, Smith PC. Hospital case payment systems in Europe. Health Care Manag Sci. 2006 Aug; 9(3): 211-3.
46. European Hospital and Healthcare Federation (HOPE), HOPE - HOPE questionnaire on DRG and hospital financing, First analysis of the results, in print, 65pp. Available from: http://www.hope.be/
47. Oostenbrink JB, Rutten FF. Cost assessment and price setting of inpatient care in The Netherlands. The DBC case-mix system. Health Care Manag Sci. 2006 Aug; 9(3): 287-94.
48. Folpmers M, de Bruijn J. Honorering in de zorg Overgang naar de nieuwe landelijke DBC-systematiek, MCA Tijdschrift voor Organisaties in Control. 2004; 18-25.
49. Department of Health NHS. Costing Manual; May 2005. 188pp. Available from: www.ic.nhs.uk/casemix
50. Epstein D, Mason A. Costs and prices for inpatient care in England: mirror twins or distant cousins? Health Care Manag Sci. 2006 Aug; 9(3): 233-42.
51. Fattore G, Torbica A. Inpatient reimbursement system in Italy: how do tariffs relate to costs? Health Care Manag Sci. 2006 Aug; 9(3): 251-8.
52. Ankjaer-Jensen A, Rosling P, Bilde L. Variable prospective financing in the Danish hospital sector and the development of a Danish case-mix system. Health Care Manag Sci. 2006 Aug; 9(3):259-68.
53. Molgaard E. Calculation of nursing costs in relation to the DRG-system, Viborg County; 2000. 84pp.
54. Reinhardt UE. The pricing of U.S. hospital services: chaos behind a veil of secrecy. Health Affairs. 2006; 25: 57-69.
55. Berkowitz E. Medicare and Medicaid: The past as prologue. Healthcare Financing Review. winter 2005-2006; 27(2): 11-23.
56. Ballard KA, Gray RF, Knauf RA, Uppal P. Measuring variations in nursing care per DRG. Nursing Management. 1993; 24(4): 33-36.
57. Welton JM. Paying for nursing care in hospitals. Am J Nurs. 2006 Nov;106(11): 67-9.
58. Centers for Medicare & Medicaid Services. Medicare Program. Proposed Changes to the Hospital Inpatient Prospective Payment Systems and Fiscal Year 2007 Rates. 2006. Available from: http://www.cms.hhs.gov/AcuteInpatientPPS/downloads/cms1488p.pdf
59. Diers D, Bozzo J. Using administrative data for practice and management. Nurs Econ. 1999 Jul-Aug; 17(4): 233-7.
60. American Hospital Association. National Uniform Billing Committee. Available from: http://www.nubc.org/
61. Welton JM, Fischer MH, Degrace S, Zone-Smith L. Nursing intensity billing. J Nurs Adm. 2006 Apr; 36(4): 181-8.
62. Welton JM, Unruh L, Halloran EJ. Nurse staffing, nursing intensity, staff mix, and direct nursing care costs across Massachusetts hospitals. J Nurs Adm. 2006 Sep; 36(9): 416-25.
63. Bellanger MM, Tardif L. Accounting and reimbursement schemes for inpatient care in France. Health Care Manag Sci. 2006 Aug; 9(3): 295-305.
64. Buisson G. DOCUMENT DE TRAVAIL, Redressements du Programme de Médicalisation des Systèmes d'Informations PMSI. France: MINISTÈRE DE L'EMPLOI, DU TRAVAIL ET DE LA COHÉSION SOCIALE ; 2005. 28pp.
65. Dubois – Fresney C. Le Résumé de soins infirmiers. Malouic; 1997. 109pp.

66. Schreyogg J, Tiemann O, Busse R. Cost accounting to determine prices: how well do prices reflect costs in the German DRG-system? Health Care Manag Sci. 2006 Aug; 9(3): 269-79.

67. Deutsche Krankenhausgesellschaft (DKG). KALKULATION VON FALLKOSTEN. Handbuch zur Anwendung in Krankenhäusern; 2002. 243pp.

68. Victorian Government Department of Human Services. Public hospitals and mental health services. Policy and funding guidelines 2006–07. Melbourne Victoria; 2006. 158pp. Available from: www.health.vic.gov.au/pfg

69. DHBNZ SFG. Casemix Cost Weights Project Group, New Zealand Casemix Framework For Publicly Funded Hospitals including WIES11B Methodology and Casemix Purchase Unit Allocation Recommended for the 2006/2007 Financial Year. 2007. 87pp. Available from: http://www.nzhis.govt.nz/

70. Commonwealth Department of Health and Ageing. National Hospital Cost Data Collection Cost Report Round 9 (2004-2005) AR-DRGv5.0. September 2006. 159pp. Available from: www.health.gov.au/internet/wcms/Publishing.nsf/Content/casemix-1

71. Australian Government Department of Health and Ageing. National Hospital Cost Data Collection Hospital Reference Manual Round 10 (2005-06) For the 2005-06 Financial Year. September 2006. 138pp. Available from: http://www.ag.gov.au/cca

72. Hovenga E, Hindmarsh C. Queensland Health – PAIS Validation Study. Results and issues for nursing cost capture. In: Proceedings of the eight casemix conference; Australia, Sydney; 1996.

73. Picone D et al. Casemix and Nursing Quality Project, Commonwealth Department of Human Services and Health, Casemix Development Program. 1995.

74. Canadian Institute for Health Information. Exploring the 70/30 Split: How Canada's Health Care System Is Financed. 2005. 147pp. Available from: www.cihi.ca

75. Canadian Institute for Health Information. DAD RESOURCE INTENSITY WEIGHTS AND EXPECTED LENGTH OF STAY, 2003 and 2004. 2004. 203pp. Available from: www.cihi.ca

76. Canadian Institute for Health Information. Canadian MIS Database Hospital Financial Performance Indicators 1999.2000 to 2003.2004 and Preliminary 2004.2005, Methodological Notes. 2006. 27pp. Available from: www.cihi.ca

77. Canadian Institute for Health Information. Future development of information to support the management of nursing resources. 2001. 30pp. Available from: www.cihi.ca

78. Canadian Nurses Association. Measuring Nurses' workload, Nursing Now. March 2003; 1-4. Available from: http://www.cna-aiic.ca/cna/

79. APDRG Suisse. Cost-Weights et Swiss Payment Groups Version 5.1. Mai 2005. 28pp. Available from: www.apdrgsuisse.ch

80. APDRG Suisse. TAR APDRG 2006: Principes et règles de financement et de facturation par APDRG. Mai 2005. 15pp. Available from: www.apdrgsuisse.ch

81. Berthou A, Junger A. Rapport final nursing data 1998 – 2000. Lausanne: ISE; 2000.

82. Rey JC, Berthou A , Junger A, Frishhuerht B. DRGs and Nursing. Bulletin de la société d'études economiques et socials. juin 2006; 64: 155-162. Available from: www.isesuisse.ch

83. Tilquin C et al. PRN 87 - Measuring the Level of Nursing Care Required. Montréal: E.R.O.S.; 1989. 175pp. .

84. O'Brien-Pallas LL, Cockerill R, Leatt P. Different systems, different costs? An examination of the comparability of workload measurement systems. Journal of Nursing Administration. 1992; 22: 17-22.

85. Cromwell J, Price K. The sensitivity of DRG weights to variation in nursing intensity. Nursing Economic. 1988; 6(1): 18-26.

86. Sermeus W, Delesie L, Van Den Heede K. Updating the Belgian nursing minimum data set: Framework and methodology. Studies in Health Technology and Informatics. 2002; 93: 89–93.

87. Bostrom J, Sutter WN. Research utilisation: making the link to practise. Journal of Nursing Staff Development. 1993; 9: 28-34.

88. Estabrooks CA. Will evidence-based nursing practice make practice perfect? Can J Nurs Res. 1998; 30(1): 15-36.

89. Thompson C. Nurses' use of research information in clinical decision making: a descriptive and analytical study. Final report. London: NHS R&D, NCC SDO; 2000. 202pp.

90. Lohr KN. Rating the strength of scientific evidence: relevance for quality improvement programs. International Journal for Quality in Health Care. 2004; 16(1): 9-18.

91. Atkins D, Eccles M, Flottorp S, et al. Systems for grading the quality of evidence and the strength of recommendations I: Critical appraisal of existing approaches The GRADE Working Group. BMC Health Services Research. 2004; 4(38): 7pp.

92. Atkins D, Briss PA, Eccles M., et al. Systems for grading the quality of evidence and the strenght of recommendations II : Pilot study of a new system. BMC Health Services Research. 2005; 5(25): 12pp.

93. Petrisor BA, Keating J, Schemitsch E. Grading the evidence: Levels of evidence and grades of recommendation. Int J Care Injured. 2006; 37: 321-327.

94. Schünemann HJ, Fretheim A, Oxman AD. Improving the use of research evidence in guideline development: 9. Grading evidence and recommendations. Health Research Policy and Systems. 2006; 4(21): 7pp.

95. Brighton B, Bhandari M, Tornetta P, Felson D. Hierarchy of evidence: From case reports to randomized controlled trials. Clin Orthopaedics and related research. 2003; 413: 19-24.

96. Busse JW, Heetveld MJ. Critical appraisal of the orthopaedic literature: Therapeutic and economic analysis. Int J Care Injured. 2006; 37: 312-320.

97. Concato J, Shah N, Horwitz R. Randomized, controlled trials, observational studies, and the hierarchy of research designs. New Engl J Med. 2000; 342(25): 1887-1892.

98. Guyatt G, Gutterman D, Baumann MH, et al. Grading strenght of recommendations and quality of evidence in clinical guidelines: report from an American College of Chest physicians Task Force. Chest. 2006; 129(1): 174-181.

99. Committee on Redesigning Health Insurance Performance Measures, Payment, and Performance Improvement Program. To err is human: building a safer health system. ISBN: 0-309-06837-1. 2000. 312pp.

100. Starfield B. Is US health really the best in the world? JAMA. 2000; 284 (4): 483-485.

101. Weingart SN, Wilson RM, Gibberd RW, et al. Epidemiology of medical error. BMJ. 2000; 320: 774-7.

102. Kirsch M. Pay-for-performance medicine – quality or quagmire? Am J Gastroenterol. 2006; 101: 2453-55.

103. Leape LL, Berwick DM. Five years after to err is human. What have we learned? JAMA. 2005; 293: 2384-90.

104. Rowe JW. Pay-for-performance and accountability: related themes in improving health care. Ann Intern Med. 2006; 145: 695-99.

105. Roland M. Linking physicians' pay to the quality of care – A major experiment in the United Kingdom. N Engl J Med. 2004; 351: 1448-54.

106. Committee on Redesigning Health Insurance Performance Measures, Payment, and Performance Improvement Programa. Rewarding Provider Performance. ISBN: 0-309-10216-2. 2007. 280pp.

107. Cognetti DM, Reiter D. The implications of "pay-for-performance" reimbursement for Otolaryngology–Head and Neck Surgery. Otolaryngology-Head and Neck Surgery. 2006; 134: 1036-42.

108. Petersen LA, Woodard LD, Urech T, et al. Does pay-for-performance improve the quality of healthcare? Annals Intern Med. 2006; 145: 265-72.

109. Lobach DF, Kerner N. A systematic process for converting text-based guidelines into a linear algorithm for electronic implementation. Proc AMIA Symp. 2000; 507-11.

110. Colombet I, Aguirre-Junco A, Zunino S, Jaulent MC, Leneveut L, Chatellier G. Electronic implementation of guidelines in the EsPeR system: A knowledge specification method. International Journal of Medical Informatics. 2005; 74: 597-604.

111. Hurst K. Nursing by numbers. Nursing standard. 2006; 21(7): 22-25.

112. Fagerström L, Rainio A. Professional assessment of optimal nursing care intensity level: a new method of assessing personnel resources for nursing care. J Clin Nursing. 1999; 8: 369-379.

113. Colombo A, Solberg B, Vanderhoeft E, Ramsay G, Schouten H. Measurement of nursing care time of specific interventions on a haematology – oncology unit related to diagnostic categories. Cancer nursing. 2005; 28(6): 476-480.

114. Gran-moravec M, Hughes C. Nursing time allocation and other considerations for staffing. Nursing and Health Sciences. 2005; 7: 126-133.

115. Burke TA, McKee JR, Wilson HC, Donahue R, Batenhorst AS, Pathak DS. A comparison of time-and-motion and self-reporting methods of work measurement. J Nurs Adm. 2000; 30: 118-125.

116. Hughes M. Nursing workload: an unquantifiable entity. Journal of nursing management. 1999; 7: 317-323.

117. Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. J Adv Nursing. 2006; 53(2): 205-212.

118. Keeney S, Hasson F, Mckenna H. A critical review of the Delphi technique as a research methodology for nursing. International Journal of Nursing Studies. 2001; 38: 195-200.

119. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nursing. 2000; 32(4): 1008-1015.

120. Greatorex J, Dexter T. An accessible analytical approach for investigating what happens between the rounds of a Delphi study. J Adv Nursing. 2000; 32(4): 1016-1024.

121. Reis M, de Rijk A, Schaufeli W. Simplified Therapeutic Intervention Scoring System: The TISS-28 items--Results from a multicenter study. Critical Care Medicine. January 1996; 24(1): 64-73.

122. Dierickx H, Sermeus W. Patientenclassificatie. Zorgenbehoefte als basis van staffing. Leuven/Amersfoort: ACCO; 1985. 150pp.

123. Hadley J, Murphy L, Dunlap L, Williams M, Mcathy M. Methods for studying. Nurse staffing in a patient unit. U.S. Departement of Health, Education and Welfare; 1978. Publication nr. HRA 78-3. 226pp.

124. Benaim C, Froger J, Compan B, Pelissier J. The assessment of autonomy in elderly people. Ann Readapt Med Phys. 2005 Jul; 48(6): 336-40.

125. Mott B, Nolan J, Zarb N, Arnison V, Chan R et al. Clinical nurses knowledge of evidence based practice: constructing a framework to evaluate a multifaceted intervention for implementing EBP. Contemp Nurse. 2005; 19: 96-104.

126. Bucknall T, Copnell B, Shannon K, McKinley D. Evidence based practice: Are critical care nurses ready for it? Aust Crit Care. 2001; 14(3): 92-8.

127. Hosmer DW, Lemeshow S. Applied Logistic Regression, 2nd edition; New York: Wiley; 2000.

128. Lesaffre E, Albert A. Multiple-group logistic regression diagnostics. Applied Statistics. 1989; 38: 425–440.

129. Albert A, Harris E. Multivariate Interpretation of Clinical Laboratory Data. New York: Marcel Dekker; 1987.

130. Nagelkerke NJD. A note on a general definition of the coefficient of determination. Biometrika. 1991; 78(3): 691-692.

131. Taheri P, Butz D, Greenfield L. Length of stay has minimal impact on the cost of hospital admission. J Am Coll Surg. 2000; 191(2): 123-130.

132. Knauf R, Ballard K, Mossman P, Lichtig L. Nursing cost by DRG: nursing intensity weights. Policy , politics and nursing practice. 2006; 7(4): 281-289.

This page is left intentionally blank.

Dépôt légal : D/2007/10.273/07

KCE reports

1. Efficacité et rentabilité des thérapies de sevrage tabagique. D/2004/10.273/2.
2. Etude relative aux coûts potentiels liés à une éventuelle modification des règles du droit de la responsabilité médicale (Phase 1). D/2004/10.273/4.
3. Utilisation des antibiotiques en milieu hospitalier dans le cas de la pyélonéphrite aiguë. D/2004/10.273/6.
4. Leucoréduction. Une mesure envisageable dans le cadre de la politique nationale de sécurité des transfusions sanguines. D/2004/10.273/8.
5. Evaluation des risques préopératoires. D/2004/10.273/10.
6. Validation du rapport de la Commission d'examen du sous financement des hôpitaux. D/2004/10.273/12.
7. Recommandation nationale relative aux soins prénatals: Une base pour un itinéraire clinique de suivi de grossesses. D/2004/10.273/14.
8. Systèmes de financement des médicaments hospitaliers: étude descriptive de certains pays européens et du Canada. D/2004/10.273/16.
9. Feedback: évaluation de l'impact et des barrières à l'implémentation – Rapport de recherche: partie 1. D/2005/10.273/02.
10. Le coût des prothèses dentaires. D/2005/10.273/04.
11. Dépistage du cancer du sein. D/2005/10.273/06.
12. Etude d'une méthode de financement alternative pour le sang et les dérivés sanguins labiles dans les hôpitaux. D/2005/10.273/08.
13. Traitement endovasculaire de la sténose carotidienne. D/2005/10.273/10.
14. Variations des pratiques médicales hospitalières en cas d'infarctus aigu du myocarde en Belgique. D/2005/10.273/12
15. Evolution des dépenses de santé. D/2005/10.273/14.
16. Etude relative aux coûts potentiels liés à une éventuelle modification des règles du droit de la responsabilité médicale. Phase II : développement d'un modèle actuariel et premières estimations. D/2005/10.273/16.
17. Evaluation des montants de référence. D/2005/10.273/18.
18. Utilisation des itinéraires cliniques et guides de bonne pratique afin de déterminer de manière prospective les honoraires des médecins hospitaliers: plus facile à dire qu'à faire.. D/2005/10.273/20
19. Evaluation de l'impact d'une contribution personnelle forfaitaire sur le recours au service d'urgences. D/2005/10.273/22.
20. HTA Diagnostic Moléculaire en Belgique. D/2005/10.273/24, D/2005/10.273/26.
21. HTA Matériel de Stomie en Belgique. D/2005/10.273.28.
22. HTA Tomographie par Emission de Positrons en Belgique. D/2005/10.273/30.
23. HTA Le traitement électif endovasculaire de l'anévrysme de l'aorte abdominale (AAA). D/2005/10.273.33.
24. L'emploi des peptides natriurétiques dans l'approche diagnostique des patients présentant une suspicion de décompensation cardiaque. D/2005/10.273.35
25. Endoscopie par capsule. D2006/10.273.02.
26. Aspects médico-légaux des recommandations de bonne pratique médicale. D2006/10.273/06.
27. Qualité et organisation des soins du diabète de type 2. D2006/10.273/08.
28. Recommandations provisoires pour les évaluations pharmacoéconomiques en Belgique. D2006/10.273/11.
29. Recommandations nationales Collège d'oncologie : A. cadre général pour un manuel d'oncologie B. base scientifique pour itinéraires cliniques de diagnostic et traitement, cancer colorectal et cancer du testicule. D2006/10.273/13.
30. Inventaire des bases de données de soins de santé. D2006/10.273/15.
31. Health Technology Assessment : l'antigène prostatique spécifique (PSA) dans le dépistage du cancer de la prostate. D2006/10.273/18.
32. Feedback: évaluation de l'impact et des barrières à l'implémentation - Rapport de recherche: partie II. D2006/10.273/20.
33. Effets et coûts de la vaccination des enfants Belges au moyen du vaccin conjugué antipneumococcique. D2006/10.273/22.
34. Trastuzumab pour les stades précoces du cancer du sein. D2006/10.273/24.

35. Etude relative aux coûts potentiels liés à une éventuelle modification des règles du droit de la responsabilité médicale – Phase III : affinement des estimations. D2006/10.273/27.
36. Traitement pharmacologique et chirurgical de l'obésité. Prise en charge résidentielle des enfants sévèrement obèses en Belgique. D/2006/10.273/29.
37. Health Technology Assessment Imagerie par Résonance Magnétique. D/2006/10.273/33.
38. Dépistage du cancer du col de l'utérus et recherche du Papillomavirus humain (HPV). D/2006/10.273/36
39. Evaluation rapide de technologies émergentes s'appliquant à la colonne vertébrale : remplacement de disque intervertébral et vertébro/cyphoplastie par ballonnet. D/2006/10.273/39.
40. Etat fonctionnel du patient: un instrument potentiel pour le remboursement de la kinésithérapie en Belgique? D/2006/10.273/41.
41. Indicateurs de qualité cliniques. D/2006/10.273/44.
42. Etude des disparités de la chirurgie élective en Belgique. D/2006/10.273/46.
43. Mise à jour de recommandations de bonne pratique existantes. D/2006/10.273/49.
44. Procédure d'évaluation des dispositifs médicaux émergeants. D/2006/10.273/51.
45. HTA Dépistage du Cancer Colorectal : état des lieux scientifique et impact budgétaire pour la Belgique. D/2006/10.273/54.
46. Health Technology Assessment. Polysomnographie et monitoring à domicile des nourrissons en prévention de la mort subite. D/2006/10.273/60.
47. L'utilisation des médicaments dans les maisons de repos et les maisons de repos et de soins Belges. D/2006/10.273/62
48. Lombalgie chronique. D/2006/10.273/64.
49. Médicaments antiviraux en cas de grippe saisonnière et pandémique. Revue de littérature et recommandations de bonne pratique. D/2006/10.273/66.
50. Contributions personnelles en matière de soins de santé en Belgique. L'impact des suppléments. D/2006/10.273/69.
51. Besoin de soins chroniques des personnes âgées de 18 à 65 ans et atteintes de lésions cérébrales acquises. D/2007/10.273/02.
52. Rapid Assessment: Prévention cardiovasculaire primaire dans la pratique du médecin généraliste en Belgique. D/2007/10.273/04.
53. Financement des soins infirmiers hospitalier. D/2007/10.273/07