Universiteit
Leiden
The Netherlands

Prediction of adverse health outcomes in older patients visiting the Emergency Department: the APOP study

Gelder, J. de

Citation

Gelder, J. de. (2019, October 17). *Prediction of adverse health outcomes in older patients visiting the Emergency Department: the APOP study*. Retrieved from https://hdl.handle.net/1887/79262

Version:	Publisher's Version
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/79262

Note: To cite this publication please use the final published version (if applicable).

Cover Page

Universiteit Leiden

The following handle holds various files of this Leiden University dissertation:
http://hdl.handle.net/1887/79262

Author: Gelder, J. de
Title: Prediction of adverse health outcomes in older patients visiting the Emergency Department: the APOP study
Issue Date: 2019-10-17

Prediction of adverse health outcomes in older patients visiting the Emergency Department

the APOP study

Jelle de Gelder

Prediction of adverse health outcomes in older patients visiting the Emergency Department: the APOP study

Jelle de Gelder

Cover design: Wendy Schoneveld - www.wenzid.nl
Lay-out: Talitha Vlastuin @ Proefschrift-AIO.nl
Print: Proefschriftmaken
ISBN: 978-94-6380-513-1

Financial support by SBOH, employer of GP trainees, for the publication of this thesis is gratefully acknowledged.

SBOH
voor artsen in opleiding

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without permission in writing from the copyright owner.
Copyright © Jelle de Gelder, 2019

Prediction of adverse health outcomes in older patients visiting the Emergency Department: the APOP study

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus Prof. Mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op donderdag 17 oktober 2019
klokke 13:45 uur

door

Jelle de Gelder

Geboren te Leiderdorp
in 1987

Promotor
Prof. dr. G.J. Blauw

Co-promotores
Dr. S.P. Mooijaart
Dr. B. de Groot

Promotiecommissie
Prof. dr. O.M. Dekkers
Prof. dr. B.M. Buurman, Universiteit van Amsterdam (UvA)
Prof. dr. C.R. Carpenter, Washington University School of Medicine

Table of contents

Chapter 1	General introduction	7
Chapter 2	Prediction of mortality in acutely hospitalized older patients	15
Chapter 3	Development and validation of the APOP screener	31
Chapter 4	Validation of the ISAR-HP	49
Chapter 5	Refinement of the APOP screener	63
Chapter 6	Predictors and outcomes of revisits	81
Chapter 7	General discussion	99
Chapter 8	English summary	111
	Nederlandse samenvatting	113
	List of co-authors	115
	List of publications	117
	Curriculum vitae	119
	Dankwoord	120

CHAPTER 1

General introduction

Abstract

General introduction

Older patients experience high rates of adverse health outcomes after visiting the emergency department (ED),[1] but at the moment of acute presentation it is difficult to determine who will deteriorate. Identification of older patients at risk for adverse health outcomes with a screening instrument may be helpful to assist health care professionals to anticipate on the risk of possible deterioration and deliver care accordingly. Several screening instruments have been developed, of which the Identification of Seniors at Risk (ISAR)[2] and Triage Risk Stratification Tool (TRST)[3] are studied most. Usability of such risk stratification screening instruments is debated, due to the limitation in distinguishing low from high risk patients and because a relatively high proportion of patients is incorrectly assigned as 'high risk'.[4, 5] To date, there still is a lack of pragmatic, accurate and reliable instruments for risk stratification of older patients in the ED.[4]

The rate of adverse health outcomes is particularly high in the first three months after the ED visit. Approximately 10% will die and up to 45% will experience functional decline.[1] A comprehensive geriatric assessment (CGA) is able to identify the older patients at increased risk and consequently improve outcomes.[6] Performing a CGA in the acute setting is virtually impossible, due to the time limitation and often the condition of the patient. Therefore, another strategy is necessary in order to identify those at high risk. A two-step approach is suggested in order to reduce the incidence of adverse health outcomes.[7] First a screening instrument is needed to identify the patients at increased risk. The second step is to target interventions in patients at highest risk. Interventions which are tailored to the individuals' need and preferences, for example determined with a CGA, can help older patients to maintain independence.[6, 8]

The ED is designed for acutely ill and injured patients and is characterized by a high patient turnover, rapid triage, acute interventions and with a focus on disposition.[1, 9] Since the 1980's scientists payed increasing attention to the health care needs of the older patients[10] and now in the 21st century the need to redesign the core of the ED is high with the ageing population.[11] The last decades research mainly focussed on the construction of a profile of older patients visiting the ED,[11-13] describing their patterns of use of emergency services[1, 14] and determining post-discharge outcomes.[15-17] The results of these studies demonstrated the need for an efficient screening instrument.

Older patients have different characteristics compared to younger patients. They arrive more often by ambulance, have a twice as high hospitalization rate, their visits are more urgent and they stay longer in the ED.[10, 18] A greater number of diagnostic tests is performed in the ED, but the final diagnoses are less accurate.[1] Beyond that, the ED is not designed for vulnerable older patients. During the relatively long length of stay it is

experienced as an unfamiliar noisy environment, while laying on a thin mattress in a room with little privacy, with the physical examinations being quickly performed and often with a restriction on nutrition and drinks.[19]

Multiple factors contribute to the complexity of delivering adequate care to the older patient in the ED.[18, 20] Approximately 20-40% of this population present with impaired cognition, but this is recognised only in a third of the cases.[11, 21-23] Older patients often present with atypical symptoms, resulting in incomplete resolution of their initial complaints.[1] They have a high prevalence of comorbidities, which causes that physicians have to deal with a mixture of chronic, subacute and acute issues.[22, 24, 25] The complex older patients in the ED challenges physicians to deliver adequate (after)care for the individual patient.

Not only scientific publications highlight the difficulties in acute care for older patients. In the Netherlands, the last years press releases payed more attention to this subject, with catchy headlines as examples of all the shared concerns: 'Spoedhulp voor ouderen in het gedrang', 'SEH raakt verstopt door vergrijzing', 'Spoedeisende hulp overspoeld door oudere mensen', 'Zelfstandige ouderen zetten SEH onder druk', 'Ouderen zoeken noodopvang in ziekenhuizen', 'Alleen ketenzorg zonder schotten kan SEH redden', 'Spoedzorg in gevaar door toenemend aantal ouderen'. The Dutch ministry of Health, Welfare and Sport (VWS) is aware of the complexity of the problem and emphasizes to explore practical solutions within the existing care system. Some local initiatives are already conducted, such as a district nurse in the emergency department, additional emergency beds for vulnerable older patients and the initiating of a hospital for low complex medical care in vulnerable older patients. The intention to improve care of the older patients visiting the ED is widely present, but it remains difficult to determine where to start. In my view we have start at the very beginning of the care process at the ED, i.e. identifying those who are at the highest risk for developing adverse health outcomes. Therefore, it was the aim of the present thesis to develop a screening instrument for the ED setting to identify older patients at high risk for adverse health outcomes, in order to target them for future interventions to preserve their independency and quality of life.

Outline of the thesis

Several aspects of prognostic prediction models are described in this thesis. In **chapter 2** available retrospective clinical data is used to develop a prediction model for 90-day mortality in acutely hospitalized older patients. In **chapter 3** we aimed to develop and validate a new prediction model for functional decline and mortality in older patients visiting the ED in a prospective study. **Chapter 4** evaluates the performance of the ISAR-HP in predicting adverse health outcomes in hospitalized older patients. In **chapter 5** previous experiences and the knowledge from the other chapters is used to refine and cross-validate the existing APOP screener. **Chapter 6** studies the predictors of ED revisits and the association with functional decline and mortality.

References

1. Aminzadeh, F. and W.B. Dalziel, *Older adults in the emergency department: a systematic review of patterns of use, adverse outcomes, and effectiveness of interventions.* Ann.Emerg.Med., 2002. 39 (3) : p. 238-247.
2. McCusker, J., et al., *Detection of older people at increased risk of adverse health outcomes after an emergency visit: the ISAR screening tool.* J.Am.Geriatr.Soc., 1999. 47 (10): p. 1229-1237.
3. Meldon, S.W., et al., *A brief risk-stratification tool to predict repeat emergency department visits and hospitalizations in older patients discharged from the emergency department.* Acad Emerg Med, 2003. 10 (3): p. 224-32.
4. Carpenter, C.R., et al., *Risk factors and screening instruments to predict adverse outcomes for undifferentiated older emergency department patients: a systematic review and meta-analysis.* Acad Emerg Med, 2015. 22 (1): p. 1-21.
5. Yao, J.L., et al., *A systematic review of the identification of seniors at risk (ISAR) tool for the prediction of adverse outcome in elderly patients seen in the emergency department.* Int J Clin Exp Med, 2015. 8 (4): p. 4778-86.
6. Ellis, G., T. Marshall, and C. Ritchie, *Comprehensive geriatric assessment in the emergency department.* Clin Interv Aging, 2014. (9): p. 2033-43.
7. McCusker, J., et al., *Rapid two-stage emergency department intervention for seniors: impact on continuity of care.* Acad.Emerg.Med., 2003. 10 (3): p. 233-243.
8. Beswick, A.D., et al., *Complex interventions to improve physical function and maintain independent living in elderly people: a systematic review and meta-analysis.* Lancet, 2008. 371 (9614): p. 725-35.
9. Adams, J.G. and L.W. Gerson, *A new model for emergency care of geriatric patients.* Acad Emerg Med, 2003. 10 (3): p. 271-4.
10. Lowenstein, S.R., et al., *Care of the elderly in the emergency department.* Ann Emerg Med, 1986. 15 (5): p. 528-35.
11. Samaras, N., et al., *Older patients in the emergency department: a review.* Ann.Emerg.Med., 2010. 56 (3): p. 261-269.
12. McCusker, J., et al., *Determinants of emergency department visits by older adults: a systematic review.* Acad Emerg Med, 2003. 10 (12): p. 1362-70.
13. Gray, L.C., et al., *Profiles of older patients in the emergency department: findings from the interRAI Multinational Emergency Department Study.* Ann Emerg Med, 2013. 62 (5): p. 467-74.
14. Gruneir, A., M.J. Silver, and P.A. Rochon, *Emergency department use by older adults: a literature review on trends, appropriateness, and consequences of unmet health care needs.* Med Care Res Rev, 2011. 68 (2): p. 131-55.
15. Sager, M.A., et al., *Functional outcomes of acute medical illness and hospitalization in older persons.* Arch Intern Med, 1996. 156 (6): p. 645-52.
16. Hastings, S.N., et al., *Adverse health outcomes after discharge from the emergency department--incidence and risk factors in a veteran population.* J Gen Intern Med, 2007. 22 (11): p. 1527-31.

17. de Saint-Hubert, M., et al., *Risk factors predicting later functional decline in older hospitalized patients.* Acta Clin.Belg., 2009. 64 (3): p. 187-194.
18. Singal, B.M., et al., *Geriatric patient emergency visits. Part I: Comparison of visits by geriatric and younger patients.* Ann Emerg Med, 1992. 21 (7): p. 802-7.
19. Hwang, U. and R.S. Morrison, *The geriatric emergency department.* J Am Geriatr Soc, 2007. 55 (11): p. 1873-6.
20. Lucke, J.A., et al., *Early prediction of hospital admission for emergency department patients: a comparison between patients younger or older than 70 years.* Emerg Med J, 2018. 35 (1): p. 18-27.
21. Litovitz, G.L., et al., *Recognition of psychological and cognitive impairments in the emergency department.* Am J Emerg Med, 1985. 3 (5): p. 400-2.
22. Salvi, F., et al., *The elderly in the emergency department: a critical review of problems and solutions.* Intern.Emerg.Med., 2007. 2 (4): p. 292-301.
23. Schofield, I., et al., *Screening for cognitive impairment in older people attending accident and emergency using the 4-item Abbreviated Mental Test.* Eur J Emerg Med, 2010. 17 (6): p. 340-2.
24. Schellevis, F.G., et al., *Comorbidity of chronic diseases in general practice.* J Clin Epidemiol, 1993. 46 (5): p. 469-73.
25. Barnett, K., et al., *Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study.* Lancet, 2012. 380 (9836): p. 37-43.

CHAPTER 2

Prediction of mortality in acutely hospitalized older patients

Predicting mortality in acutely hospitalized older patients:
a retrospective cohort study

J. de Gelder, J.A. Lucke, N. Heim, A.J. de Craen, S.D. Lourens, E.W. Steyerberg, B. de Groot, A.J. Fogteloo, G.J. Blauw, S.P. Mooijaart

Published in Internal Emergency Medicine 2016; 11(4):587-94

Abstract

Acutely hospitalized older patients have an increased risk of mortality, but at the moment of presentation this risk is difficult to assess. Early identification of patients at high risk might increase the awareness of the physician, and enable tailored decision-making. Existing screening instruments mainly use either geriatric factors or severity of disease for prognostication. Predictive performance of these instruments is moderate, which hampers successive interventions. We conducted a retrospective cohort study among all patients aged 70 years and over who were acutely hospitalized in the Acute Medical Unit of the Leiden University Medical Center, the Netherlands in 2012. We developed a prediction model for 90-day mortality that combines vital signs and laboratory test results reflecting severity of disease with geriatric factors, represented by comorbidities and number of medications. Among 517 patients, 94 patients (18.2%) died within 90 days after admission. Six predictors of mortality were included in a model for mortality: oxygen saturation, Charlson comorbidity index, thrombocytes, urea, C-reactive protein and non-fasting glucose. The prediction model performs satisfactorily with an 0.738 (0.667-0.798). Using this model, 53% of the patients in the highest risk decile (N = 51) were deceased within 90 days. In conclusion, we are able to predict 90-day mortality in acutely hospitalized older patients using a model with directly available clinical data describing disease severity and geriatric factors. After further validation, such a model might be used in clinical decision making in older patients.

Introduction

Acute medical illness in older adults is a serious contributor to deterioration.[26] Within 90 days after hospitalization, approximately 20 % will die.[27, 28] At the time of admission it is difficult to determine who is at highest risk. Visualizing the individual risk in an early phase of hospitalization might increase the awareness of the physician, and enable tailored decision-making for the older patient, although these interventions may not primarily be aimed at reducing mortality. A high risk of mortality may reflect overall vulnerability, which preventive interventions may be aimed at, or conversely by usefully initiating palliative care.

Screening models to identify older patients at risk of mortality have been developed and evaluated.[29, 30] These models mainly use either geriatric factors[30, 31]or severity of disease.[32, 33] In these models, scores are assigned to the predictors, which lead to a total score with a cut-off point for high-risk patients. Predictive performance using the cut-off point shows relatively high sensitivity and low specificity, resulting in high numbers of false positives. As a consequence, large-scale implementation of these models in daily care hampers successive interventions.[31] A combination of routine clinical parameters, which reflect the severity of disease, in combination with geriatric factors might improve accuracy and feasibility in daily care.

In the present retrospective cohort study we developed a prediction model for 90-day mortality. We collected clinical parameters of all hospitalized older patients of the acute medical unit in 2012. Vital signs and laboratory results reflect severity of disease with geriatric factors, represented by comorbidities and number of medications. We selected variables that are available in a very early phase of hospitalization to enable in-hospital interventions.

Methods

Study design and setting

We performed a retrospective follow-up study among all patients aged 70 years and over who were acutely hospitalized on the acute medical unit (AMU) of the Leiden University Medical Center (LUMC), the Netherlands in 2012. Any following individual admission in the study period, independent of the reason, and patients with palliative care who were expected to die in a few days were excluded. The AMU is a 13-bed ward particularly focussed on acute admissions, mainly from the Emergency Department. The population is characterized by hemodynamically stable patients in the fields of internal medicine, surgery, pulmonary diseases and gastroenterology. The medical ethics committee of the LUMC waived the necessity for formal approval of the present study, as all data were available from standard care.

Predictors

We selected potential predictors of 90-day mortality from the clinical parameters available at the moment of hospitalization on the AMU. These parameters reflect severity of disease, including vital signs and laboratory results, and underlying level of vulnerability, including comorbidity and number of medications used at home. A predictor was eligible if it fulfilled the following criteria: (1) it was available in the medical records for retrospective analysis; (2) available to the physician within 24 h after admission and (3) assumed to have a relationship to the outcome based on clinical reasoning by three medical doctors and (4) was already being measured routinely to enhance in future implementation, with a maximum of 15 % missing values of each predictor. Multiple imputation techniques were used to compute the missing predictors.[34] First measured vital signs after hospitalization and first known in-hospital laboratory results were extracted from the electronic patient records (Chipsoft-EZIS®, version 5.2, 2006–2014). Existing comorbidities and medications used at home were obtained manually from the patient records, where medication was reported as part of routine clinical practice. Usually, the physician will first ask the patient at the moment of hospitalization for comorbidities and medication use. If necessary, the information will be verified with the general practitioner or pharmacy.

Vital signs were assessed by the nurse directly after admission and consisted of systolic and diastolic blood pressure, heart rate, respiratory rate, oxygen saturation and body temperature. First known in-hospital laboratory results within 24 h after presentation were extracted and consisted of: sodium (mmol/L), potassium (mmol/L), urea (mmol/L), eGFR (estimated glomular filtration rate, calculated by the modification of diet in renal disease (MDRD) equation, ml/min/1.73 m^2), leukocytes ($\times 10^9$/L), thrombocytes ($\times 10^9$/L), C-reactive protein (mg/L), non-fasted glucose (mmol/L) and haemoglobin (mmol/L). Comorbidity was evaluated with the Charlson comorbidity index (CCI). It incorporates weighted scores

for 19 medical conditions, increasing from 1 to 6 with severity. The CCI is a frequently used instrument by researchers, and has been validated to predict 1-year mortality.[35, 36] The number of different medications at home was recorded according to their pharmacological sub classification. Medications of the same subgroup count as one drug, and combined medications of two different pharmacological sub-classifications were considered as two different drugs. Topical and 'as required' medications were excluded because of the unreliable registration rate of the physicians and the absence of information whether the patient actually used it. If recorded in the medical records, over-the-counter medications were included when patients used it on regular base.

Outcome

The primary endpoint was mortality within 90 days after hospital admission. Mortality dates were assessed from the Dutch municipality records.

Coding predictors

We aimed to develop a model with a high positive predictive value (PPV) to enable targeted interventions, and therefore the model should have high specificity. Because a model based on a risk score derived from clinical cut-off values is easier to implement in clinical practice than a model based on computations with continuous variables, we started to dichotomize continuous predictors by using the ranges of clinical reference categories. A stricter clinically relevant cut-off point was chosen in cases when specificity on 90-day mortality was lower than fifty percent. However, dichotomizing may lead to loss of information, reduction in power and uncertainly in defining the optimal cutpoint.[37] In a sensitivity analysis, we repeated the same analyses with preservation of continuous predictors to compare discriminative performance.

Statistical/data analysis

Descriptive baseline characteristics were expressed in percentages, means with standard deviations and medians with interquartile ranges. A Kaplan–Meier curve was used to show cumulative mortality of patients after admission. Binary logistic regression techniques for both dichotomised and continuous data were used for uni- and multivariable analysis on mortality (no censoring occurred). The prediction model was derived via backward elimination with Akaike's Information Criterion for candidate predictors. With this technique the most significant predictors remain in the model, while "noise" is reduced by eliminating predictors that are not statistically significant. Discriminative performance of the different models was assessed with the area under the receiver operating characteristic curve (AUC).The model was internal validated with 500 bootstrap samples, where we repeated the backward elimination procedure. With this method internal validity is tested by drawing 500 different population samples from the original sample. The test reflects how robust the findings are when slight changes are made to the population, and was

presented with the internal validated AUC. The formula $1/(1 + \exp^{(-\text{linear predictor})})$was applied to determine the individual risk on 90-day mortality.[38] Performance of the final model is shown with the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio and negative likelihood ratio.

In clinical practice, cut-off points are often used, e.g., to interpret laboratory results. In prediction dichotomizing continuous predictors is arguable, because of a loss of information.[37] When using cut-off points, laboratory results just outside the reference range are considered the same risk as more extreme results, which diminishes the power of the model. In a sensitivity analyses we treated predictors as continuous variables to compare discrimination. First outlying observations were truncated to 5 and 95 % by means of winsorization.[38] Second restricted cubic spline techniques with three knots were applied to continuous predictors in a binary regression model, and discrimination was analysed by calculation the AUC. The level of significance was set at $P < 0.05$. Statistical analyses were performed using IBM SPSS Statistics package (version 20) and R version 3.1.1.

Results

In 2012, 606 older patients were acutely hospitalized to the acute medical unit (AMU) of our hospital. By excluding 86 subsequent admissions and 3 moribund patients, a final cohort of 517 patients was available for final analysis.

The baseline characteristics of the cohort are described in Table 1. The mean age was 78.3 years, 269 (52.0 %) patients were male, 467 (90.3 %) were admitted via the Emergency Department and 367 (71.0 %) were primary treated under the responsibility of one of the medical specialities (internal medicine, surgery or pulmonary diseases). Mean laboratory results were within the normal range or slightly below or above these thresholds. The median number of comorbidities was 2 (IQR 1-4), and median number of medications used at home was 7 (IQR 4-11).

Supplemental Table 1 gives an overview of categories of the dichotomization process. Missing values, up to 11 % in thrombocytes, were imputed. Most reference ranges reflect clinical normal ranges, except that we chose different rounded cut-offs for systolic blood pressure (<200 mmHg), urea (<15.0 mmol/L) and c-reactive protein (CRP, <100 mg/L), leukocytes ($<13 \times 10^9$/L), eGFR (>30 ml/min/1.73 m^2) and haemoglobin (>6.5 mmol/L for females and >7.5 mmol/L for males) to reflect more specific measures of disease severity.

Table 1: Baseline characteristics of the study population

Characteristic	N = 517
Male, *n (%)*	269 (52.0%)
Age, *mean (SD)*	78.3 (6.2)
Admitted from, *n (%)*	
Emergency department	467 (90.3)
Outpatient clinic	42 (8.1)
Other	8 (1.6)
Clinical specialism, *n (%)*	
Internal medicine	367 (71.0%)
Surgery	74 (14.3%)
Lung diseases	34 (6.6%)
Other	42 (8.1%)
Severity of disease Vital parameters[1]	
Oxygen saturation (%), *median (IQR)*	98 (96-99)
Systolic blood pressure (mm Hg), *mean (SD)*	132.9 (23.3)
Diastolic blood pressure (mm Hg), *mean (SD)*	67.6 (13.9)
Heart rate (/min), *mean (SD)*	83.2 (16.6)
Laboratory results	
Sodium (mmol/L), *mean (SD)*	138.6 (5.4)
Potassium (mmol/L), *mean (SD)*	4.3 (0.7)
Urea (mmol/L), *median (IQR)*	9.4 (6.7-14.7)
eGFR (ml/min/1,73m2), *mean (SD)*	64.8 (34.5)
Leukocytes (x 109/L), *mean (SD)*	11.3 (5.3)
Thrombocytes (x 109/L), *mean (SD)*	241 (119)
C-reactive protein (mg/L), *median (IQR)*	41 (8-110)
Non-fasted glucose (mmol/L), *mean (SD)*	8.1 (3.6)
Haemoglobin (mmol/L), *mean (SD)*	7.6 (1.5)
Geriatric factors	
Charlson comorbidity index, *median (IQR)*[2]	2 (1-4)
Number of medications, *median (IQR)*	7 (4-11)

eGFR = Estimated Glomerular Filtration Rate, SD = Standard Deviation, IQR = Inter Quartile Range
1) Respiratory rate and body temperature were excluded from further analysis, because the measurement was not performed or noted in >50%.
2) Incorporates weighted scores for 19 medical conditions, higher scores indicating worse history of disease.

Table 2: Univariate associations and the performance of single predictors of 90-day mortality in acutely hospitalized older patients

		Univariate		Performance				
	Number (%)[1]	**OR**	**95% CI**	**Sens**	**Spec**	**PPV**	**NPV**	**AUC**
Age (per 5 years increase)	-	1.05	0.88-1.25	-	-	-	-	0.52
Male	269 (52%)	1.38	0.88-2.17	0.59	0.49	0.20	0.84	0.54
Saturation	108 (21%)	2.21	1.35-3.63	0.33	0.82	0.29	0.85	0.57
Systolic blood pressure	39 (8%)	1.39	0.64-3.03	0.10	0.93	0.23	0.82	0.51
Diastolic blood pressure	206 (40%)	1.21	0.77-1.90	0.44	0.61	0.20	0.83	0.52
Heart rate	136 (26%)	1.68	1.04-2.71	0.35	0.76	0.24	0.84	0.55
Charlson comorbidity index[2]	91 (18%)	3.64	2.20-6.03	0.36	0.87	0.37	0.86	0.61
Number of medications	153 (30%)	1.53	0.96-2.45	0.37	0.72	0.23	0.84	0.55
Thrombocytes	139 (27%)	2.16	1.35-3.46	0.40	0.76	0.27	0.85	0.58
Urea	126 (24%)	2.44	1.52-3.92	0.39	0.79	0.29	0.85	0.59
Leukocytes	169 (33%)	1.43	0.90-2.27	0.39	0.69	0.22	0.84	0.54
Sodium	162 (31%)	1.16	0.72-1.87	0.34	0.69	0.20	0.83	0.52
Potassium	134 (26%)	1.73	1.07-2.78	0.35	0.76	0.25	0.84	0.56
Haemoglobin	161 (31%)	1.57	0.98-2.49	0.39	0.71	0.23	0.84	0.55
C-reactive protein	141 (27%)	1.77	1.11-2.85	0.37	0.75	0.25	0.84	0.56
Non-fasted glucose	361 (70%)	0.51	0.32-0.81	0.43	0.73	0.26	0.85	0.58
eGFR	76 (15%)	0.48	0.28-0.84	0.23	0.87	0.29	0.84	0.55

eGFR = Estimated Glomerular Filtration Rate

1) Number represents number of patients outside reference category.

2) Incorporates weighted scores for 19 medical conditions, higher scores indicating worse history of disease.

After 90 days, 94 patients (18.2 %) had died (supplemental Figure. 1). In Table 2, results of the univariate analyses and the performance of all individual predictors are shown. Oxygen saturation, heart rate, Charlson comorbidity index (CCI), thrombocytes, urea, potassium and CRP outside the reference range are statistically significantly associated with 90-day mortality. In contrast, non-fasted glucose and creatinine clearance outside the reference range prove to have a protective effect. Age and gender show no association with 90-day mortality. Best performance of a single variable is the CCI. A score of 5 or higher ($N = 91$) yields a positive predictive value (PPV) of 0.37 and area under the curve (AUC) of 0.61.

Table 3: Multivariate and final model of predictors of 90-day mortality in acute hospitalized older patients

	Multivariate		**Final model**			
	OR	**95% CI**	**β**	**OR**	**95% CI**	**P-value**
Age (per 5 years)	1.20	0.97-1.47				
Male	1.34	0.79-2.25				
Saturation	2.32	1.34-4.03	0.862	2.37	1.39-4.05	0.002
Systolic blood pressure	0.96	0.38-2.45				
Diastolic blood pressure	1.02	0.60-1.72				
Heart rate	1.58	0.93-2.70				
Charlson comorbidity index[1]	3.45	1.95-6.10	1.201	3.32	1.94-5.70	<0.001
Number of drugs	1.43	0.79-2.58				
Thrombocytes	2.10	1.24-3.56	0.774	2.17	1.30-3.62	0.003
Urea	1.90	0.98-3.66	0.706	2.03	1.21-3.41	0.008
Leukocytes	1.24	0.73-2.11				
Sodium	1.08	0.64-1.82				
Potassium	1.38	0.80-2.40				
Haemoglobin	1.06	0.62-1.81				
C-reactive protein	1.65	0.96-2.81	0.588	1.80	1.08-2.99	0.023
Non-fasted glucose	0.44	0.26-0.73	-0.791	0.45	0.28-0.75	0.002
eGFR	1.24	0.57-2.69				
Intercept			-2.127			
AUC (95% CI)			0.738 (0.678-0.798)			
Internal validated AUC			0.724			

eGFR = Estimated Glomerular Filtration Rate, AUC = Area Under Curve
Internal validated AUC = the obtained AUC after bootstrapping with backward selection.
1) Incorporates weighted scores for 19 medical conditions, higher scores indicating worse history of disease.

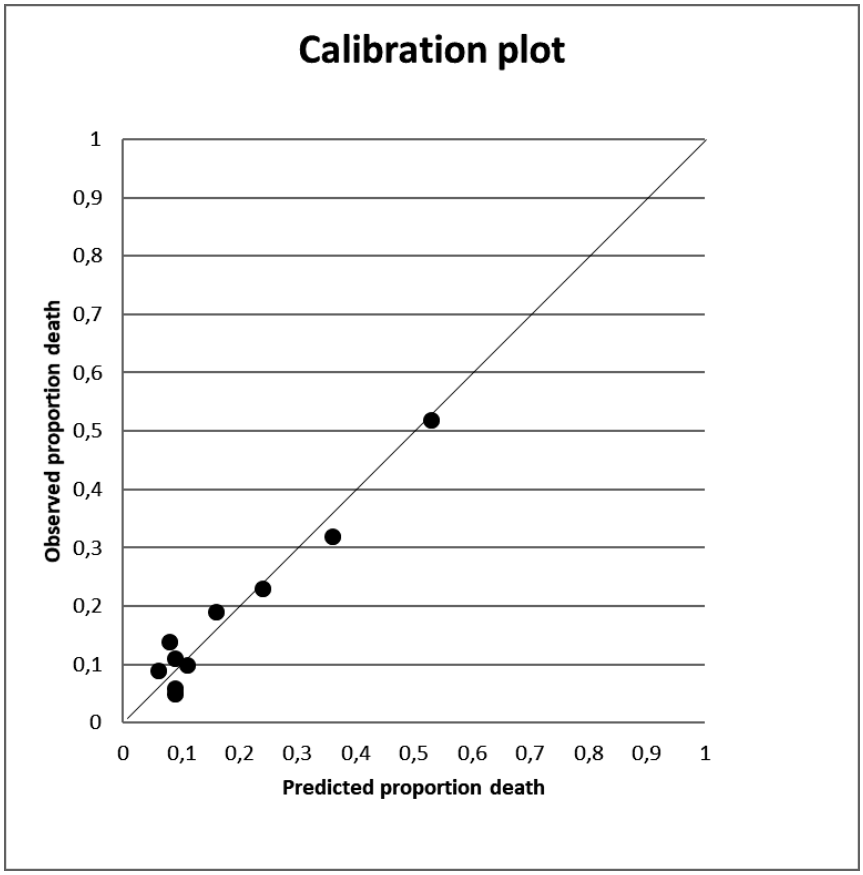

Figure 1: Comparison of observed and predicted 90 days mortality for acutely hospitalized patients into 10 equal groups

Results of the multivariable and final model are displayed in Table 3. A backward selection procedure results in a model of six predictors including oxygen saturation, CCI, thrombocytes, urea, CRP and non-fasted glucose. The area under the curve (AUC) is 0.738 (95 %CI 0.967–0.798) and decreases to 0.724 after internal validation.

By applying the formula 1/(1 + exp(–(–2.127 + 0.862 × 'saturation' + 1.201 × 'CCI' + 0.774 × 'thrombocytes' + 0.706 × 'urea' + 0.588 × 'CRP' + –0.791 × 'non-fasted glucose'))) the individual risk of 90-day mortality in acutely hospitalized older patients was calculated. Figure 1 shows the calibration plot of the final model. Over the whole range predicted probabilities are in line with the observed, with more spread in the higher risk groups. In Table 4 we calculated predictive performance of the 30, 20 and 10 % patients at highest risk to provide information about clinical usefulness. Positive likelihood ratio's range from 2.70 to 5.06. The positive predicting value of the 51 patients (10 %) at highest risk is 0.53 and implies that 53 % die within 90 days after admission.

As a sensitivity analysis we repeated analyses for the multivariate and final model with continuous data. Accuracy is comparable in both multivariate and final model. The AUC for continuous data is 0.771 (95 %CI 0.717–0.825) and after dichotomization 0.758 (95 %CI 0.702–0.815) in the multivariate model and 0.736 (95 %CI 0.677–0.795) and 0.738 (95 %CI 0.678–0.798) in the final model (data not shown).

Table 4: Performance of predicted high risk deciles in older hospitalized patients

	Number of patients	Sens	Spec	PPV	NPV	LR+	LR-
30% high risk	160	0.64	0.76	0.38	0.90	2.70	0.47
20% high risk	106	0.50	0.86	0.44	0.89	3.58	0.58
10% high risk	51	0.29	0.94	0.53	0.86	5.06	0.76

Sens = sensitivity, Spec = specificity, PPV = positive predicting value, NPV = negative predicting value, LR+ = positive likelihood ratio, LR- = negative likelihood ratio

Discussion

In the present study, we developed a prediction model for 90-day mortality in acutely hospitalized older patients using routinely collected clinical parameters describing disease severity and geriatric factors. With this model we are able to identify a high-risk group with an average 53 % risk of mortality within 90 days after admission compared to the baseline risk of 18.2 %.

We developed and internally validated a prediction model for 90-day mortality in acutely hospitalized older patients using six routinely collected clinical parameters and with adding age and gender. Underlying vulnerability of older patients is reflected in the Charlson comorbidity index (CCI) and severity of disease in oxygen saturation, thrombocytes, urea, C-reactive protein and non-fasted glucose. CCI was developed to predict mortality in medical patients,[35] and was recently validated in acutely hospitalized older adults to predict both short- and long-term mortality.[36] Models using vital signs have also been previously studied to predict mortality. The acute physiology and chronic health evaluation (APACHE II) is a severity of disease classification system developed to predict in-hospital mortality in intensive care unit patients of all ages.[39] The APACHE II comprises a combination of vital parameters and different laboratory results. The simple clinical score (SCS) is a prediction model for 30-day mortality in acutely admitted patients,[32] and consists of 16 parameters, including vital parameters and presentation signs such as new stroke, coma and abnormal ECG. The Modified Early Warning Score (MEWS) was originally designed for the ED setting to identify medical patients at risk of catastrophic deterioration, and was subsequently validated for prediction of in-hospital mortality in hospitalized patients.[33] The MEWS incorporates five vital signs: systolic

blood pressure, pulse rate, respiratory rate, body temperature and level of consciousness. The aforementioned models are well validated and are used in practice, but share the disadvantage that prognostic accuracy among older patients is modest with relatively low positive predicting values. An explanation might be the use of (bed-side) scores with a cut-off point, instead of using individual risk scores. Or it could be the use of either severity of disease characteristics or geriatric factors in the prediction model. Another explanation could be that prediction models were developed in a more severely ill population of all ages, with the consequence that results were neither representative nor tailored for these older patients.[40, 41] Unexpected findings is the positiveness of abnormal thrombocytes and urea. To our knowledge, these measurements are not used in other comparable prediction models. Validity of this might be explained by the possible over-representation of patients with low thrombocytes being treated with chemotherapy or high urea caused by dehydration or kidney failure. Another unexpected finding is the protective value of creatinine clearance <30 (ml/min/1.73m^2) on 90-day mortality (OR 0.48, 95 % CI 0.28–0.84) in the univariate analysis. A possible explanation could be that the hospital is a centre for patients requiring dialysis and kidney transplantation. These patient groups are hospitalized more readily, with possible less severe acute medical conditions. However, in the multivariate model and by using creatinine clearance as a continuous variable the association is lost, indicating that it could also be caused by outliers. Taken together, we show that combining parameters reflecting severity of disease and geriatric factors results in an prediction model capable of predicting 90-day mortality in acutely hospitalized older patients.

We made several choices in developing our model, in order to be ready for clinical implementation. First, we used routinely available clinical parameters. Candidate predictors were known within 24 h after admission. Second, we used a formula instead of a bed-side scorecard. Health care professionals do not have to calculate risks by hand, preferable the formula should be integrated in the electronic patients records or be available in a smartphone application. By providing an individual risk for each patient, the consequences of the screening can vary. Depending on the local resource availability a hospital can implement interventions. As an example, comprehensive geriatric assessment could be performed in all older patients with a risk of 30 % or higher, which includes extra attention for patient preferences, treatment goals and possible palliative trajectories. In another hospital the advice for the treating physician could be to take into account both the individual risk score and the condition of the patient in decision making, without standardized interventions. Third, we aimed for a model with a high specificity, resulting in a high positive predictive value (PPV) in relation with the baseline risk. The PPV will give the clinician a robust feeling and may therefore be more relevant than a high AUC with modest predictive values. The identification senior at risk (ISAR) is also a prediction tool, originally developed to identify older patients in the ED at increased risk of adverse health

outcomes, a composite outcome of functional decline and mortality.[2] The ISAR is a widely used tool in the ED,[5] and validated among 667 acute hospitalized older adults for prediction of adverse outcomes, including mortality. After 90 days of follow-up 5 % had died, with 6 % of the patients assigned high risk deceased within 90 days, indicating a low positive predictive value. The negative predictive value (NPV) for 90-day mortality was 0.97, which means that 97 % of the patients not at risk were still alive after 90 days. These results imply that the ISAR in this setting is more suitable to rule out patients at high risk, whereas our model is tailored to identify older patients at high risk for mortality with a PPV of 0.53 in the highest risk group. Identifying of patients at low risk ("rule-out") may be a very sensible strategy in its own right. However, our aim is to specifically identify patients at the highest risk because these are the patients we want to follow-up with intervention, and we want to aim our limited clinical resources to only those at the highest risk.

Our study has several limitations. First, we studied retrospective data, and therefore the number of available predictors and related outcomes were limited. Ideally, predictors such as cognition, functional status and outcomes such as functional decline, and readmissions should also be used, but these were not available in this retrospective study. Second, the fact that we found some unexpected results further stresses the need for external validation, as it is impossible to distinguish whether these findings are specific to our cohort, chance finding or reproducible in other cohorts. Strengths of the present study are that our prediction model is based on routinely measured and directly available candidate predictors. This enhances convenient future implementation in an early phase of presentation. We used clinical cut-off points to reflect clinical practice and relate to the awareness of the physician. From a methodological point of view using continuous variables is preferable, but is harder to relate to clinical practice. Nevertheless, accuracy of our model is equally well when dichotomized or with continuous variables. Another strength is the high specificity of the developed model. This specificity ensures the development of interventions that are aimed at a relatively small group of patients at high risk of a negative event. Such tools are of importance in the emergency medicine setting, allowing physicians in EDs and Acute Wards to make informed decisions on diagnostic and therapeutic strategies in older patients and the implementation of measures to prevent poor outcome.

In conclusion, we developed a prediction model on 90-day mortality in acutely hospitalized older patients. We used a combination of predictors containing information about severity of disease and geriatric factors and calculated individual risk scores. Currently, we are conducting a large multicentre prospective follow-up study among acutely presenting older patients in both the ED and wards, including more candidate predictors and outcomes.

References

1. Clegg, A., et al., *Frailty in elderly people.* Lancet, 2013. 381 (9868): p. 752-762.
2. Covinsky, K.E., et al., *Loss of independence in activities of daily living in older adults hospitalized with medical illnesses: increased vulnerability with age.* J Am Geriatr Soc, 2003. 51 (4): p. 451-8.
3. Smolin, B., et al., *Predicting mortality of elderly patients acutely admitted to the Department of Internal Medicine.* Int J Clin Pract, 2014.
4. Campbell, S.E., et al., *A systematic literature review of factors affecting outcome in older medical patients admitted to hospital.* Age Ageing, 2004. 33 (2): p. 110-5.
5. Wou, F., et al., *The predictive properties of frailty-rating scales in the acute medical unit.* Age Ageing, 2013. 42 (6): p. 776-81.
6. Edmans, J., et al., *The Identification of Seniors at Risk (ISAR) score to predict clinical outcomes and health service costs in older people discharged from UK acute medical units.* Age Ageing, 2013. 42 (6): p. 747-53.
7. Kellett, J. and B. Deane, *The Simple Clinical Score predicts mortality for 30 days after admission to an acute medical unit.* QJM, 2006. 99 (11): p. 771-81.
8. M. Cei, C.B., N. Mumoli, *In-hospital mortality and morbidity of elderly medical patients can be predicted at admission by the Modified Early Warning Score: a prospective study.* Int J Clin Pract., 2009. 63 (4): p. 591-5.
9. Janssen, K.J., et al., *Dealing with missing predictor values when applying clinical prediction models.* Clin Chem, 2009. 55 (5): p. 994-1001.
10. Charlson, M.E., et al., *A new method of classifying prognostic comorbidity in longitudinal studies: development and validation.* J Chronic Dis, 1987. 40 (5): p. 373-83.
11. Frenkel, W.J., et al., *Validation of the Charlson Comorbidity Index in acutely hospitalized elderly adults: a prospective cohort study.* J Am Geriatr Soc, 2014. 62 (2): p. 342-6.
12. Royston, P., D.G. Altman, and W. Sauerbrei, *Dichotomizing continuous predictors in multiple regression: a bad idea.* Stat Med, 2006. 25 (1): p. 127-41.
13. Steyerberg, E.W., *Clinical prediction models: a practical approach to development, validation, and updating.*, ed. M.G.K.K.J.S.A.T. W.Wong. 2009, New York: Springer.
14. Knaus, W.A., et al., *APACHE II: a severity of disease classification system.* Crit Care Med, 1985. 13 (10): p. 818-29.
15. Broekhuizen, K., et al., *Characteristics of randomized controlled trials designed for elderly: a systematic review.* PLoS One, 2015. 10 (5): p. e0126709.
16. Mooijaart, S.P., et al., *Evidence-based medicine in older patients: how can we do better?* Neth J Med, 2015. 73 (5): p. 211-8.
17. McCusker, J., et al., *Detection of older people at increased risk of adverse health outcomes after an emergency visit: the ISAR screening tool.* J.Am.Geriatr.Soc., 1999. 47 (10): p. 1229-1237.
18. Yao, J.L., et al., *A systematic review of the identification of seniors at risk (ISAR) tool for the prediction of adverse outcome in elderly patients seen in the emergency department.* Int J Clin Exp Med, 2015. 8 (4): p. 4778-86.

CHAPTER 3

Development and validation of the APOP screener

Predicting adverse health outcomes in older emergency department patients: the APOP study

Authors: J. de Gelder, J.A. Lucke, B. de Groot, A.J. Fogteloo, S. Anten, K. Mesri, E.W. Steyerberg, C. Heringhaus, G.J. Blauw, S.P. Mooijaart

Published in The Netherlands Journal Medicine 2016; 74(8):342-352.

Abstract

BACKGROUND: Older patients experience high rates of adverse outcomes after an emergency department (ED) visit. Early identification of those at high risk could guide preventive interventions and tailored treatment decisions, but available models perform poorly in discriminating those at highest risk. The present study aims to develop and validate a prediction model for functional decline and mortality in older patients presenting to the ED.

METHODS: A prospective follow-up study in patients aged >/= 70, attending the EDs of the LUMC, the Netherlands (derivation) and Alrijne Hospital, the Netherlands (validation) was conducted. A baseline assessment was performed and the main outcome, a composite of functional decline and mortality, was obtained after 90 days of follow-up.

RESULTS: In total 751 patients were enrolled in the Leiden University Medical Center of whom 230 patients (30.6%) experienced the composite outcome and 71 patients (9.5%) died. The final model for the composite outcome resulted in an area under the curve (AUC) of 0.73 (95% CI 0.67-0.77) and was experienced in 69% of the patients at highest risk. For mortality the AUC was 0.79 (95% CI 0.73-0.85) and 36% of the patients at highest risk died. External validation in 881 patients of Alrijne Hospital showed an AUC of 0.71 (95% CI 0.67-0.75) for the composite outcome and 0.67 (95% CI 0.60-0.73) for mortality.

CONCLUSION: We successfully developed and validated prediction models for 90-day composite outcome and 90-day mortality in older emergency patients. The benefits for patient management by implementing these models with preventive interventions have to be investigated.

Introduction

Older patients presenting to the emergency department (ED) experience high rates of adverse outcomes,[1] but they form a heterogeneous group and it is unknown who is at highest risk. The incidence of adverse outcomes is particularly high after three months, with a mortality rate about 10% and increased functional dependence between 10-45%. [1] Early identification of those at highest risk gives an opportunity to guide preventive interventions and informed treatment decisions.[8]

Current models use either severity of disease or existing geriatric vulnerability for prediction. The Modified Early Warning Score (MEWS) is an indicator of disease severity and showed to be valuable in predicting worse in-hospital outcomes in older patients. [42] However, prognostication of MEWS for long-term outcomes in older adults is unknown. The Identification of Seniors At Risk (ISAR)[2] and Triage Risk Stratification Tool (TRST)[3] focus on existing geriatric vulnerabilities, such as functional and cognitive impairment, to predict adverse health outcomes. Neither of these tools accurately identify high-risk patients,[4, 5] while that vulnerable group of patients benefits most from an increased level of attention.

We conducted a prospective follow-up study among all older patients who visited the EDs in the region of Leiden. The aim was to develop and validate a prediction model for adverse health outcomes in older emergency patients. To reflect the condition of the patient, both demographics and severity of disease indicators and existing geriatric vulnerability were taken into account.

Methods

Study design and setting

We performed a prospective follow-up study in the EDs of two hospitals in the region of Leiden, the Netherlands. The Leiden University Medical Center (LUMC, derivation cohort) is an academic hospital with a level 1 trauma centre and Alrijne Hospital (location Leiderdorp, validation cohort) is a peripheral hospital with a level 2 trauma centre. We considered all patients eligible who fulfilled the inclusion and exclusion criteria. The inclusion criterion was patients aged ≥ 70 presenting for the first time to the ED in the study period. The exclusion criteria were patients who were triaged with highest urgency (code red), who we were not able to approach due to an unstable medical condition, lack of permission of the nurse or physician to enter the room for any reason or due to impaired mental status without an authorised relative to provide informed consent. Also a language barrier and patients who left the waiting room were not eligible. Patients were enrolled 7 days a week for 12 weeks with 24-hour coverage in the LUMC and 12-hour coverage (10.00 am to 10.00 pm) in Alrijne Hospital. Written informed consent was obtained before inclusion. The medical ethics committee of the LUMC and Alrijne waived the necessity for formal approval of the present study, as the study closely follows routine care.

Organisation of emergency care in the Netherlands

Basic health insurance is mandatory in the Netherlands and covers the care from general practitioners (GPs), hospital care and specialist care. Emergency care is provided by GPs and in EDs. Patients in need of immediate care can contact the GP, the GP out-of-hours service, call for an ambulance or go to an ED by themselves. Depending on the urgency, patients are expected to contact or visit the GP first. As with the structure of the LUMC and Alrijne, an increasing number of GP out-of-hours centres are integrated close to the ED to avoid unnecessary ED visits.[43] In EDs a triage nurse will prioritise patients first based on the severity of their condition; then the patient can either be directed to the emergency room or the waiting room. In the Leiden region, the LUMC and Alrijne Hospital are the only two EDs, together servicing an unselected catchment area of 400,000 inhabitants of all ages. In both EDs there are no special rooms or trajectories for older patients. Two patient groups bypass the ED and were therefore impossible to include: 1) Older patients with a myocardial infarction in the ambulance who were directly sent to the catheterisation room and 2) older patients with a CVA and eligible for thrombolytic therapy underwent a brief primary assessment in the ED and were then sent to the neurology ward after a CT scan.

Data collection

We included patients in the LUMC from September to November 2014 and in Alrijne Hospital from March to June 2015. In both hospitals teams of medical students were present at the ED from 10.00 am until 10.00 pm to enrol patients, and in the LUMC the ED staff were responsible for inclusion from 10.00 pm until 10.00 am. Before the start of the inclusion period, the medical students and ED staff of the LUMC attended training sessions to guarantee convergence on conducting the questionnaires. The ideal moment for conducting the questionnaires turned out to be 30-45 minutes after arrival of the patient to the ED. At that moment the patient had spoken to the physician and was waiting for lab results or further analysis. The questionnaire took 5-10 minutes to complete. A representative was permitted to answer questions when the patient was unable to provide answers, with the exception of the cognition and self-reported quality of life questions. Questions were collected on a tablet computer and sent directly to a secured database. Additional medical data were extracted automatically from the medical records, verified manually and added to the database.

Baseline

At baseline, data on three domains were assessed: demographics, severity of disease indicators and geriatric measurements. Demographics consisted of age, gender, living arrangements and level of education. A low level of education was defined as elementary school, basic education as community college, middle education as secondary education and high education as higher vocational training or university. Severity of disease indicators consisted of characteristics related to the ED visit: way of arrival, triage category by the Manchester Triage System,[44] fall-related ED visit, indication to measure vital signs and indication to perform a blood test. Whether the visit was fall related was obtained by asking the question: Is the reason for presentation related to a fall? Indication to measure vital signs or laboratory tests was scored positively when, at the moment of presentation, vital signs needed to be measured or a laboratory test was ordered based on the Manchester Triage System and local protocols. Geriatric measurements consisted of the number of different medications mentioned by the patient, history of diagnosed dementia reported by patient or proxy, current use of a walking device, the Identification of Seniors at Risk (ISAR)[2] screening tool, the Six Item Cognitive Impairment test (6CIT) [45] and the Katz Index of Activities of Daily Living (ADL)[46] questionnaire. The ISAR was developed for patients aged ≥ 65 and aims to predict the risk of adverse health outcomes six months after the ED visit. The ISAR consists of six dichotomous questions and scores range from 0 to 6 with higher scores denoting higher risk. The 6CIT is a short cognition test and was validated in a Dutch population against the Mini Mental State Examination (MMSE)[47] with a score on the 6CIT of ≥ 11 indicating cognitive impairment (MMSE < 24).[48] Six questions lead to a score ranging from 0 to 28 with higher scores indicating more cognitive impairment. The Katz ADL indicates functional status two weeks before

presentation to the ED to eliminate possible effects of the acute illness and consists of six dichotomous questions on dependence in bathing, dressing, toileting, transfer, eating and incontinence. Scores range from 0 to 6 with higher scores an indication of more dependency.

Outcomes

The main outcome of the study was composite outcome, a composite of functional decline or mortality at 90-day follow-up. Functional decline was defined as at least one point increase in the Katz ADL score or new institutionalisation, defined as a higher level of assisted living at 90 days after ED visit. We analysed 90-day mortality separately. Mortality can be seen as the ultimate decline and might then be taken together with functional decline. On the other hand, the intervention strategy could differ for patients at high risk for mortality. For that reason we developed a separate prediction model for 90-day mortality. A model solely for functional decline is not feasible. Excluding deceased patients would imply that the model is only applicable in patients who will not die within a certain period, which we do not know at the moment of presentation. Three months after the ED visit the patient was contacted by telephone. In case of no response after three attempts on three consecutive days, the GP was contacted to verify the phone number and living status. Finally a letter with the follow-up questions was sent to patients who had not moved to a higher level of assisted living and who were alive according to the information from the GP. Data concerning mortality were derived from the municipal records.

Statistical analysis

Baseline characteristics are presented as mean with standard deviation (SD) in case of normal distribution, median with interquartile range (IQR) in case of skewed distribution or as numbers with percentages (%). Adequate statistical power for obtaining good predictions requires a minimum of 10 events per candidate predictor.[38] This rule was followed for the composite outcome. To reduce the number of candidate predictors, the most relevant questions from the questionnaires (Katz ADL, ISAR, 6CIT) were pre-selected. Single questions with the highest R-square values on the entire questionnaire score were selected and added to the list of candidate predictors for development of the prediction model. Missing predictors were imputed via single imputation techniques.[49] The prediction model was derived via backward elimination with Akaike's Information Criterion (equivalent to $p < 0.157$ for predictors with 1 df). With this technique the least contributing candidate predictors are deleted until the deterioration in model fit is too large. Discrimination of the models was assessed with the area under the receiver operating characteristic curve (AUC). Internal validation was conducted with a 500 bootstrap sample procedure, where we repeated the backward elimination procedure in each bootstrap sample to estimate the optimism-corrected performance that is expected if the derived prediction model is applied in other datasets. The internal validation procedure also

provided a shrinkage factor to adjust the estimated regression coefficients for overfitting.[50] The adjusted regression equation provides predictions for new individuals. It was validated in the Alrijne patients.[51] Calibration of the model, which reflects how well predicted and observed outcomes agree, was examined by using the adjusted regression equation. Calibration was examined graphically with calibration plots, with a goodness of fit test (Hosmer and Lemeshow test[52]). The formula $1/(1+e^{(-\text{linear predictor})})$ was applied with the adjusted regression equation to determine the individual risks of experiencing the outcome. Performance of the model for the patients with the highest 30%, 20% and 10% predicted risk was evaluated according to sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio and negative likelihood ratio, with 95% confidence intervals. To compare our model performance with the existing six-item ISAR questionnaire, predictive performance of the ISAR on different cut-off points was also calculated. The level of statistical significance was set at $p < 0.05$. Statistical analyses were performed using IBM SPSS Statistics package (version 20) and R version 3.1.1.

Results

In the three-month inclusion period a total of 995 older patients presented to the ED of the LUMC. Of these, 19 patients were excluded due to a language barrier or leaving the waiting room. Another 92 patients could not be approached due to their medical condition, resulting in 884 eligible patients. Of these, 65 patients were missed for inclusion and 68 patients refused informed consent, which led to a study population of 751 patients (85% of eligible patients). Similarly, 881 patients were included in Alrijne Hospital (figure 1).

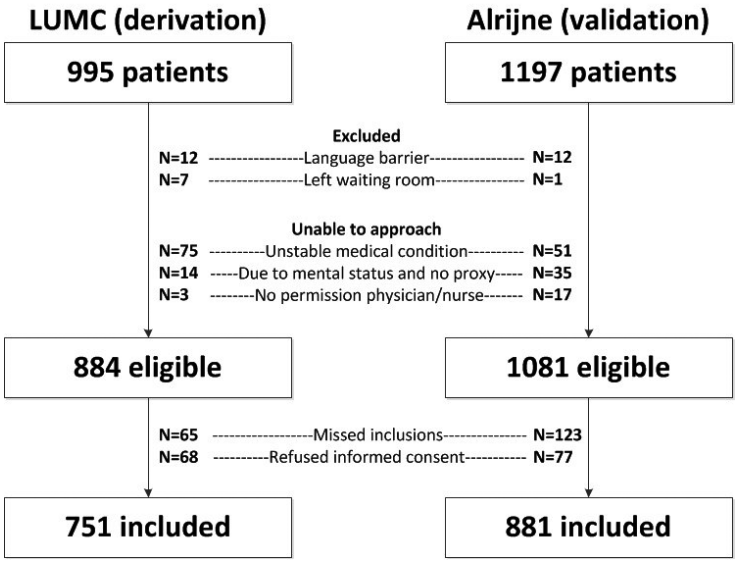

Figure 1: Flowchart of the study population

Table 1: Baseline characteristics of older emergency patients at LUMC and Alrijne

Characteristics	LUMC (N = 751)*	Alrijne (N = 881)#
Demographics		
Age (years), median (IQR)	78 (74-83)	80 (75-84)
Female	389 (51.8%)	454 (51.5%)
Living in residential care or nursing home, *n (%)*	63 (8.4%)	69 (7.8%)
High education, *n (%)*	155 (20.8%)	164 (18.6%)
Severity of disease indicators		
Arrival by ambulance, n(%)	405 (53.9%)	432 (49.0%)
Triage category, n(%)		
Standard (Green)	159 (21.2%)	353 (40.1%)
Urgent (Yellow)	391 (52.1%)	470 (53.3%)
Very urgent (Orange)	201 (26.8%)	58 (6.6%)
Fall related visit, n(%)	211 (28.1%)	192 (21.8%)
Indication to perform vital sign measurement(s), n(%)	661 (88.0%)	776 (88.1%)
Indication to perform blood test, n(%)	603 (80.3%)	749 (85.0%)
Geriatric measurements		
Number of different medications, median (IQR)	5 (3-8)	5 (3-8)
Use of walking device, n(%)	302 (40.4%)	378 (42.9%)
Katz ADL score, median (IQR)[1]	0 (0-1)	0 (0-1)
ISAR score, median (IQR)[2]	2 (1-3)	2 (1-3)
History of dementia, n(%)	34 (4.5%)	42 (4.8%)
6CIT score, median (IQR)[3]	4 (2-8.5)	4 (0-8)

N = number, IQR = Interquartile range
* LUMC data incomplete for education (N = 746), use of walking device (N = 747), Katz ADL score (N = 745), ISAR score (N = 748) and 6CIT score (N = 697).
# Alrijne data incomplete for education (N = 878), use of walking device (N = 878), Katz ADL score (N = 859), ISAR score (N = 872), 6CIT (N = 791)
[1] Higher scores indicating higher dependency (0-6)
[2] Higher scores indicating higher risk on functional decline (0-6)
[3] Higher scores indicating worse cognition (0-28)

The median age of LUMC patients was 78 years (IQR 74-83) and 80 years (IQR 75-84) in Alrijne patients (table 1); 405 (53.9%) patients of the LUMC arrived by ambulance and 201 (26.1%) were triaged as 'very urgent'. In Alrijne 432 (49.0%) arrived by ambulance and 58 (6.6%) were triaged as 'very urgent'. In both hospitals the majority of the older patients were independent, with a median Katz ADL score of 0 (IQR 0-1).

Table 2: Final model for 90-day composite adverse outcome and 90 day mortality in older patients

	Composite adverse outcome		Mortality	
	beta	***OR (95% CI)***	***beta***	***OR (95% CI)***
Demographics				
Age (per 5 years increase)	0.293	1.34 (1.16-1.54)	0.462	1.59 (1.28-1.97)
Female			-1.397	0.25 (0.14-0.45)
Living in residential care or nursing home			0.730	2.08 (0.94-4.58)
High education				
Severity of disease indicators				
Arrival by ambulance	0.477	1.61 (1.14-2.28)		
Triage category				
Standard (Green)				
Urgent (Yellow)				
Very urgent (Orange)				
Fall related ED visit			-0.627	0.53 (0.26-1.10)
Indication of vital measurement(s)				
Indication of blood test(s)			1.254	3.50 (1.24-9.93)
Geriatric measurements				
Number of different medications	0.044	1.05 (1.00-1.09)		
Use of walking device				
Need help bathing/showering	0.665	1.94 (1.24-3.05)		
Need help dressing			1.281	3.60 (1.97-6.58)
Hospital admission in the past 6 months	0.404	1.50 (1.03-2.17)		
Need help prior to ED visit	0.716	2.05 (1.38-3.03)		
History of dementia	-0.794	0.45 (0.20-1.01)		
Disorientated in time				
Intercept	-6.557		-10.538	
AUC (95%CI) of final model	0.73 (0.69-0.77)		0.79 (0.73-0.85)	
AUC (95%CI) in Alrijne patients	0.71 (0.67-0.75)		0.67 (0.60-0.73)	

ED = Emergency Department, AUC = Area Under the Receiver Operator Curve, CI = confidence interval
The internal validation procedure resulted in two model equations.
1: 90 day composite adverse outcome: 1/(1+exp(-(-6.557 + 0.293 x '(age/5)' + 0.477 x 'arrival by ambulance' + 0.044 x 'number of medications' + 0.665 x 'need help bathing or showering' + 0.404 x 'hospital admission in the past six months' + 0.716 x 'need help prior to ED visit' + -0.716 x 'history of dementia')))
2: 90 day mortality: 1/(1+exp(-(-10.538 + 0.462 x '(age/5)' + -1.397 x 'female gender' + 0.730 x 'living in residential care/nursing home' + -0.627 x 'fall related ED visit' + 1.254 x 'indication of blood test' + 1.281 x 'need help dressing')))

In total 230 LUMC patients (30.6%) experienced the composite outcome within 90 days of follow-up and 71 patients (9.5%) died. In Alrijne Hospital 247 patients (28.0%) had the composite outcome and 84 (9.5%) died.

Details on the univariate and multivariable analyses on both outcomes can be found in supplemental tables 1 and 2. The final model for the composite outcome combined age, arrival by ambulance, number of different medications, help needed with bathing or showering, hospital admission in the past six months, help needed at home on a regular base and history of dementia (table 2). Ninety-day mortality could be best predicted by combining information of age, gender, living arrangements, a fall prior to ED visit, indication for blood tests and needing help in dressing. Accuracy of the final models was fair to good, with in the derivation cohort an area under the curve (AUC) of 0.73 (95% CI 0.69-0.77) for the composite outcome and of 0.79 (95% CI 0.73-0.85) for mortality. External validation in Alrijne patients showed an AUC of 0.71 (95% CI 0.67-0.75) for the composite outcome and 0.67 (95% CI 0.60-0.73) for mortality. The formula of the original final models to calculate the individual risk can be found in the legend of table 2.

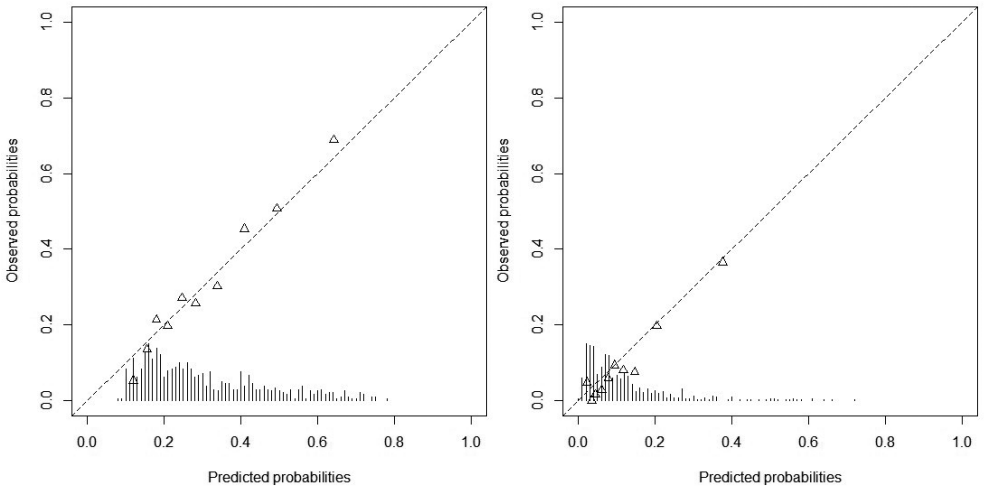

Figure 2:

A: Calibration plot at internal validation of 90-day composite outcomes with a Hosmer and Lemeshow goodness-of-fit p-value of 0.77. The vertical lines represent the relative frequency distribution of predicted probabilities
B: Calibration plot at internal validation of 90-day mortality with a Hosmer and Lemeshow goodness-of-fit p-value of 0.19. The vertical lines represent the relative frequency distribution of predicted probabilities

Calibration of the predicted probabilities was satisfactory (figure 2), with all Hosmer and Lemeshow goodness-of-fit p-values above 0.05. A stricter limit to assign patients at high risk increased specificity, PPV and the positive likelihood ratio (table 3). The PPV ranged from 0.55 (95% CI 0.48-0.61) to 0.69 (95% CI 0.57-0.79) for the composite outcome and from 0.21 (95% CI 0.16-0.27) to 0.36 (95% CI 0.26-0.49) for mortality, depending on the threshold chosen. This implies that in the highest risk group 69% of the patients experienced the composite outcome and 36% died.

The conventional cut-off of 2 points or higher for the ISAR resulted in a PPV of 0.40 (95% CI 0.36-0.45) for composite outcome and 0.13 (95% CI 0.10-0.16) for mortality (supplemental table 3). Raising the ISAR cut-off to a score of ≥ 4 points to define patients as high risk yielded a PPV of 0.50 (95% CI 0.40-0.60) for the composite outcome and 0.18 (95% CI 0.12-0.26) for mortality.

Table 3: Predictive performance of final prediction model for 90-day composite adverse outcome and 90-day mortality

	Number of patients	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% CI)	LR- (95% CI)
30% at highest risk							
Composite adverse outcome	225	0.53 (0.47-0.60)	0.80 (0.77-0.84)	0.55 (0.48-0.61)	0.80 (0.76-0.83)	2.73 (2.21-3.38)	0.58 (0.50-0.67)
Mortality	225	0.66 (0.54-0.77)	0.74 (0.70-0.77)	0.21 (0.16-0.27)	0.95 (0.93-0.97)	2.53 (2.05-3.12)	0.46 (0.33-0.63)
20% at highest risk							
Composite adverse outcome	150	0.39 (0.32-0.45)	0.88 (0.85-0.91)	0.59 (0.51-0.67)	0.77 (0.73-0.80)	3.30 (2.48-4.40)	0.69 (0.62-0.77)
Mortality	152	0.59 (0.47-0.70)	0.84 (0.81-0.86)	0.28 (0.21-0.36)	0.95 (0.93-0.97)	3.66 (2.82-4.73)	0.49 (0.37-0.65)
10% at highest risk							
Composite adverse outcome	75	0.23 (0.17-0.29)	0.96 (0.93-0.97)	0.69 (0.57-0.79)	0.74 (0.70-0.77)	5.12 (3.21-8.16)	0.81 (0.75-0.87)
Mortality	74	0.38 (0.27-0.50)	0.93 (0.91-0.95)	0.36 (0.26-0.49)	0.94 (0.91-0.95)	5.50 (3.67-8.25)	0.67 (0.55-0.80)

Sens = sensitivity, Spec = specificity, PPV = positive predictive value; NPV = negative predictive value; LR+ = positive likelihood ratio; LR- = negative likelihood ratio, CI = confidence intervals

Predictive performance was calculated by applying the results of internal validation procedure.

Discussion

New externally validated prediction models were presented for older emergency patients by using a combination of demographics, severity of disease indicators and geriatric vulnerability. Performance of the models was satisfactory, with good accuracy and high PPVs.

The predictors used in our models have previously been shown to be predictive of negative health outcomes in other models. The Identification Seniors at Risk (ISAR)[2] tool and Triage Risk Screening Tool (TRST)[3] were developed for older patients at the ED. The ISAR is suitable for all older patients, whereas the TRST was developed for those discharged home. Both tools include predominately geriatric vulnerabilities, such as functional and cognitive impairment, and are validated for prediction of negative health outcomes, including functional decline and mortality.[2, 53, 54] Scoring systems for disease severity are also used to predict negative health outcomes of which the Acute Physiology and Chronic Health Evaluation (APACHE II)[39] and Early Warning Score (MEWS)[42] are well known. APACHE II is available online and predicts mortality in intensive care unit patients by using an algorithm consisting of 12 physiological and two disease-related variables. The MEWS weighs the severity of five physiological parameters to identify patients at risk of clinical deterioration and can be used as a bedside evaluation instrument to predict mortality and admission in ED patients.[55] The MEWS and APACHE II scores were developed for prediction of worse in-hospital outcomes, whereas the prognostic capabilities in the longer term are unknown, especially in the older population. Recently we showed that directly available clinical data describing disease severity and geriatric vulnerability can be used for prediction in hospitalised older patients.[56] The present study also selected predictors reflecting the acute condition of the older patient visiting the ED and developed prediction models with high specificity and high PPVs.

As shown in table 2, demographics and severity of disease indicators are important for predicting mortality and geriatric measurements for predicting the composite outcome. It is arguable that those patients with functional impairment at baseline have a higher risk of further decline, and this stresses the importance of obtaining these measurements of functional capacity in combination with the other parameters for accurate prediction. We showed that history of dementia decreased the risk of the composite outcome. It was an unexpected finding and it may be caused by a larger proportion of patients with dementia living in an institution, and thus better protected from poor outcome, or the group of patients with dementia were less severely ill but referred to the ED sooner. Alternatively, it could be a chance finding in a small group of patients. Another notable finding is that a comparable percentage of patients in both hospitals arrived by ambulance, while the patients in LUMC are triaged more urgently (table 1). We do not have an explanation for this

finding, since there are many reasons to arrive by ambulance and both hospitals use the Manchester Triage System. However, we showed that patients who arrive by ambulance are at increased risk to experience the composite outcome. Expected or unexpected, the final prediction models have to be tested in a different population or setting to support general applicability. External validation of both models in the Alrijne patients resulted in a comparable discrimination for the composite outcome and a decrease in AUC of 0.12 for mortality. It is difficult to explain the reason for this decrease in mortality. The fact that the inclusion timeframe was different between hospitals (24 hours in the LUMC vs. 10 am to 10 pm in Alrijne Hospital) is unlikely to have influenced the results substantially, as there were only a very limited number of patients included during the night, and endpoints did not differ between those included during the 'daytime' vs. those included at 'night'. More likely, it could be minor differences in the study population, in ED protocols or parameters which we did not or cannot measure.

Predictive performance of a comparable model, the ISAR,[2] was analysed in the same study population (supplemental table 3). The performance was characterised with high sensitivities and low specificities, resulting in relatively low PPVs and high NPVs. As a consequence ISAR is more useful to 'rule out' patients at high risk, where our models target patients at highest risk. Prediction of individual risk scores on multiple outcomes, as shown with the composite outcome and mortality enable emergency physicians to guide preventive interventions and tailored treatment decisions. As an example, for patients with a predicted risk for the composite adverse outcome of 50% to 65%, safety procedures could be applied, whereas a predicted risk of 65% or higher can lead to more intensive interventions. On one hand standardised interventions should be administered, such as nursing these patients in a comfortable bed and informing the general practitioner. On the other hand, the predicted risk could support the physician in deciding to start physiotherapy or in making an outpatient appointment to prevent deterioration. If the risk of 90-day mortality is also high, this could be an argument to spend more time on diagnostic and therapeutic shared decision making and advanced care planning.

To date, there is no standard screening program for older adults in the ED. This could be due to the low proportion of evidence-based studies designed for the elderly,[40, 41] specifically due to the low number of clinical impact studies.[57] The ultimate goal is to introduce a new generalised prediction tool, suitable for all older emergency patients and to design and test effectiveness of different interventions. Such a model should consist of patient-related parameters rather than organisation-dependent factors, such as the indication to perform measurements. The present model is accurate for older patients in Western Europe. We are planning such external validation studies, which will show whether the model needs to be updated to specific settings. The algorithm can simply be

integrated in the electronic patient record to incorporate screening into routine care or be used as an application as developed on the website: http://screener.apop.eu.

One of the limitations in the current study is the lack of baseline data on potentially important determinants such as malnutrition, depression and instrumental ADL functioning. Since time is scarce in the acute setting we had to limit the number of questions, instead of performing a comprehensive geriatric assessment. A second limitation is the low proportion of deceased patients within 90 days of follow-up. As a consequence, power for prediction of 90-day mortality was low. The major strength is the unselected representative study population. We included 85% of the eligible older patients 24/7 during 12 weeks. A second strength is the fact that demographics, severity of disease and geriatric vulnerability of the patient were taken into account as a reflection of the condition of the patient.

In conclusion, we successfully developed and validated prediction models for 90-day composite outcome and 90-day mortality in older emergency patients. The benefits for patients by implementing these models with preventive interventions have to be further investigated.

References

1. Aminzadeh, F. and W.B. Dalziel, *Older adults in the emergency department: a systematic review of patterns of use, adverse outcomes, and effectiveness of interventions.* Ann.Emerg.Med., 2002. 39 (3): p. 238-247.
2. Samaras, N., et al., *Older patients in the emergency department: a review.* Ann.Emerg.Med., 2010. 56 (3): p. 261-269.
3. Cei, M., C. Bartolomei, and N. Mumoli, *In-hospital mortality and morbidity of elderly medical patients can be predicted at admission by the Modified Early Warning Score: a prospective study.* Int J Clin Pract, 2009. 63 (4): p. 591-5.
4. McCusker, J., et al., *Detection of older people at increased risk of adverse health outcomes after an emergency visit: the ISAR screening tool.* J.Am.Geriatr.Soc., 1999. 47 (10): p. 1229-1237.
5. Meldon, S.W., et al., *A brief risk-stratification tool to predict repeat emergency department visits and hospitalizations in older patients discharged from the emergency department.* Acad Emerg Med, 2003. 10 (3): p. 224-32.
6. Carpenter, C.R., et al., *Risk factors and screening instruments to predict adverse outcomes for undifferentiated older emergency department patients: a systematic review and meta-analysis.* Acad Emerg Med, 2015. 22 (1): p. 1-21.
7. Yao, J.L., et al., *A systematic review of the identification of seniors at risk (ISAR) tool for the prediction of adverse outcome in elderly patients seen in the emergency department.* Int J Clin Exp Med, 2015. 8 (4): p. 4778-86.
8. Schafer, W., et al., *The Netherlands: health system review.* Health Syst Transit, 2010. 12 (1): p. v-xxvii, 1-228.
9. Mackway-Jones, K., *Manchester Triage Group.* Emergency Triage, 1997.
10. Katzman R, B.T., Fuld P, Peck A, Schechter R, Schimmel H., *Validation of a short Orientation-Memory-Concentration Test of cognitive impairment.* Am J Psychiatry., 1983. Jun (140 (6)): p. 734-9.
11. Katz, S., et al., *Studies of Illness in the Aged. The Index of Adl: A Standardized Measure of Biological and Psychosocial Function.* JAMA, 1963. (185): p. 914-9.
12. Folstein, M.F., S.E. Folstein, and P.R. McHugh, *"Mini-mental state". A practical method for grading the cognitive state of patients for the clinician.* J.Psychiatr.Res., 1975. 12 (3): p. 189-198.
13. Tuijl, J.P., et al., *Screening for cognitive impairment in older general hospital patients: comparison of the Six-Item Cognitive Impairment Test with the Mini-Mental State Examination.* Int J Geriatr Psychiatry, 2012. 27 (7): p. 755-62.
14. Steyerberg, E.W., *Clinical prediction models: a practical approach to development, validation, and updating.*, ed. M.G.K.K.J.S.A.T. W.Wong. 2009, New York: Springer.
15. Donders, A.R., et al., *Review: a gentle introduction to imputation of missing values.* J Clin Epidemiol, 2006. 59 (10): p. 1087-91.
16. Van Houwelingen, J.C. and S. Le Cessie, *Predictive value of statistical models.* Stat Med, 1990. 9 (11): p. 1303-25.

17. Moons, K.G., et al., *Risk prediction models: I. Development, internal validation, and assessing the incremental value of a new (bio)marker.* Heart, 2012. 98 (9): p. 683-90.
18. Hosmer DW , L.S., *Applied logistic regression*. 2nd ed. 2000, New York: Wiley.
19. Hustey, F.M., et al., *A brief risk stratification tool to predict functional decline in older adults discharged from emergency departments.* J.Am.Geriatr.Soc., 2007. 55 (8): p. 1269-1274.
20. Buurman, B.M., et al., *Risk for poor outcomes in older patients discharged from an emergency department: feasibility of four screening instruments.* Eur J Emerg Med, 2011. 18 (4): p. 215-20.
21. Knaus, W.A., et al., *APACHE II: a severity of disease classification system.* Crit Care Med, 1985. 13 (10): p. 818-29.
22. Bulut, M., et al., *The comparison of modified early warning score with rapid emergency medicine score: a prospective multicentre observational cohort study on medical and surgical patients presenting to emergency department.* Emerg Med J, 2014. 31 (6): p. 476-81.
23. de Gelder, J., et al., *Predicting mortality in acutely hospitalized older patients: a retrospective cohort study.* Intern Emerg Med, 2016. 11 (4): p. 587-94.
24. Broekhuizen, K., et al., *Characteristics of randomized controlled trials designed for elderly: a systematic review.* PLoS One, 2015. 10 (5): p. e0126709.
25. Mooijaart, S.P., et al., *Evidence-based medicine in older patients: how can we do better?* Neth J Med, 2015. 73 (5): p. 211-8.
26. Steyerberg, E.W., et al., *Prognosis Research Strategy (PROGRESS) 3: prognostic model research.* PLoS Med, 2013. 10 (2): p. e1001381.

CHAPTER 4

Validation of the ISAR-HP

Optimising the ISAR-HP to screen efficiently
for functional decline in older patients

J. de Gelder, E. Haenen, J.A. Lucke, H.G. Klop, L.C. Blomaard, R.A.J. Smit,
K. Mesri, B. de Groot, A.J. Fogteloo, S. Anten, G.J. Blauw, S.P. Mooijaart

Published in The Netherlands Journal Medicine 2017; 75(9):379-385.

Abstract

INTRODUCTION: The Identification of Seniors At Risk-Hospitalised Patients (ISAR-HP) has recently been included in guidelines as a frailty indicator to identify patients for comprehensive geriatric assessment. Previous studies showed that the conventional cut-off score incorrectly classifies a high percentage of patients as high risk. We aimed to optimise the predictive value of ISAR-HP by using different cut-offs in older acutely hospitalised patients.

METHODS: A prospective follow-up study was performed in two Dutch hospitals. Acutely hospitalised patients aged ≥ 70 years were included. Demographics, illness severity parameters, geriatric measurements and the ISAR-HP scores were obtained at baseline. The primary outcome was a combined end point of functional decline or mortality during 90-day follow-up.

RESULTS: In total 765 acutely hospitalised older patients were included, with a median age of 79 years, of whom 276 (36.1%) experienced functional decline or mortality. The conventional ISAR-HP cut-off of ≥ 2 assigned 432/765 patients (56.5%) as high risk, with a positive predictive value (PPV) of 0.49 (95%CI 0.45-0.54) and a negative predictive value of 0.81 (95%CI 0.76-0.85). Thus, 51% of those whom the ISAR-HP denoted as high risk did not experience the outcome of interest. Raising the cut-off to ≥ 4 assigned 205/765 patients (26.8%) as high risk, with a marginally increased PPV to 0.55 (95%CI 0.48-0.62).

CONCLUSION: The ISAR-HP with the conventional cut-off of ≥ 2 incorrectly identifies a large group of patients at high risk for functional decline or mortality and raising the cut-off to 4 only marginally improved performance. Caution is warranted to ensure efficient screening and follow-up interventions.

Introduction

In the Netherlands, at the suggestion of both the Health Care Inspectorate (IGZ) and insurance companies, the Identification of Seniors At Risk – Hospitalised Patients (ISAR-HP) screening instrument is currently being promoted for use as a frailty indicator, for example in older patients with an indication for colon surgery.[58] Comprehensive geriatric assessment (CGA) is subsequently advised for 'frail' patients in order to prevent functional decline. Identification of patients at high risk for functional decline is essential to ensure that interventions are targeted effective at those who will benefit most.[59]

The ISAR-HP is a recently developed screening instrument to predict 90-day functional decline in older patients who were acutely admitted to the department of internal medicine.[60] Test characteristics were reasonable with respect to discrimination (area under the receiver operating curve, AUC), but the positive predictive value was rather low. Using the conventional cut-off score of ≥ 2 classified more than half of all older patients as being at risk for functional decline.[60, 61] However, classification was incorrect for 57% of the internal medicine patients in the development cohort and 64% of older patients undergoing cardiac surgery in a validation cohort, because no functional decline was experienced.[60, 61] As a consequence it is questionable whether using intensive interventions, such as the relatively time-consuming CGA, can be cost-effective.

The aim of the present study was to evaluate the performance of the ISAR-HP in predicting adverse health outcomes in acutely hospitalised older patients in two hospitals in the Netherlands. Predictive performance was tested by using different cut-off points of the ISAR-HP for predicting functional decline or mortality.

Methods

Study design and setting

The Acutely Presenting Older Patients (APOP) study is a prospective multicentre cohort study in older patients visiting the emergency department. Data were collected in the emergency departments of the Leiden University Medical Center from September 2014 until November 2014 and the Alrijne Hospital Leiderdorp from March 2015 until May 2015. In both hospitals patients were included 7 days a week for a period of 12 weeks. The inclusion criterion of the APOP study was all patients aged 70 years and older who visited the emergency department. Exclusion criteria were red on the Manchester Triage System (i.e. patients requiring acute medical attention, such as cardiopulmonary resuscitation), an unstable medical condition, refusal to participate by the patient, an impaired mental condition of patients in the absence of a proxy to provide informed consent, and presence of a language barrier. For the current analyses, all acutely hospitalised patients of the APOP cohort with an ISAR-HP score at baseline were included. The ISAR-HP scores were calculated afterwards and not noted in the patient records, to ensure that all patients received usual care. Written informed consent was obtained from all patients. The Medical Ethics Committee of the Leiden University Medical Center and Alrijne Hospital approved the study. A more detailed description of the study design can be found in a previously published paper.[62]

Characteristics

Baseline characteristics included age, gender, living situation, level of education, clinical specialism, number of medications, history of dementia, Katz ADL score and cognitive impairment. Independent living situation represents patients living independently on their own or with others, high education was defined as higher vocational training or university, and number of medications represents the number of medications used at home as reported by the patient. Clinical specialism corresponds to the responsible specialism on the ward patients were admitted to. The cognitive status was assessed with the six-item Cognitive Impairment Test (6CIT);[45] this score ranges from 0 to 28, with a score of 11 or higher indicating moderate to severe cognitive impairment. Functionality two weeks prior to admission was evaluated by means of the Katz ADL score, which contains six items: bathing, dressing, toileting, transferring, eating and the use of incontinence material.[46]Each item is scored as independent (0 points) or dependent (1 point), with higher scores corresponding to more dependency.

ISAR-HP

The ISAR-HP is a scorecard with four yes/no questions on needing assistance on a regular basis, use of a walking device, needing assistance for travelling and having received education after the age of 14 years (*Appendix, figure 1*). [60] The score ranges from 0 to

5 with a score of 2 or more indicating a high risk of functional decline. The originally developed regression model of the ISAR-HP was: 1 / 1 + exp (– (–1.93 + 0.48 × 'pre-admission need for assistance in IADL on a regular base' + 0.81 x 'use of a walking device' + 0.57 x 'need of assistance in travelling' + 0.42 × 'no education after age 14')).

Outcomes

Originally the ISAR-HP was developed for predicting solely functional decline,[60] but at the moment of obtaining an ISAR-HP score it is impossible to distinguish patients who will not die within 90 days of follow-up from those who will. Therefore, in the present study the ISAR-HP was validated for predicting the composite outcome of functional decline or mortality within 90 days of follow-up after hospital admission. Information on functional dependency was assessed by telephone. Functional decline was defined as either an increase of at least 1 point on the Katz ADL score 90 days after hospitalisation compared with two weeks prior to admission or moving from an independent living situation to a dependent living situation. Dates of death were obtained from the Dutch municipality records. Additionally, the ISAR-HP was validated for solely functional decline, for which we used the same exclusion criteria as the development study.[60] Patients with a maximum Katz ADL score at baseline (fully dependent patients) and patients living in a nursing home at baseline were excluded, because these patients could not decline further as defined in our study. Also patients who were lost to follow-up or died within 90 days were excluded.

Statistical analysis

The baseline characteristics are presented as numbers with percentages or medians with interquartile ranges (IQR). A minimum of 100 events was considered necessary to provide sufficient statistical power for external validation.[63] Predictive performance of the ISAR-HP was assessed by examining measures of discrimination and calibration. Discrimination of the ISAR-HP score was quantified by calculating the AUC. The sensitivity, specificity, positive and negative predictive values (PPV and NPV) and positive and negative likelihood ratio were calculated for using the conventional cut-off of ≥ 2 points, but also using other thresholds of the ISAR-HP score (≥ 1, ≥ 3 and ≥ 4). Calibration of the internally validated ISAR-HP regression equation was assessed by plotting observed versus predicted probabilities, calculating calibration slope and with a goodness-of-fit test (Hosmer and Lemeshow test[52]). Data were analysed using IBM SPSS Statistics version 23 (IBM Corp, Armonk, NY) and R Statistics version 3.3.0.[64].

Results

In the APOP study, 1965 consecutive older patients visiting the emergency department of the Leiden University Medical Center or Alrijne Hospital were eligible for participation. In total 1632 patients (83.1%) were included after informed consent, of whom 771 (42.2%) were subsequently hospitalised. After exclusion of three missing and three incomplete ISAR-HP scores, the study population for the present analyses contains 765 patients. Table 1 shows the baseline characteristics of the study population. The median age was 79 years (IQR 74-84), 374 patients (48.9%) were male and 698 patients (91.2%) were living independently either on their own or with others. Most patients were admitted for the clinical specialism internal medicine (242 patients, 36.1%), cardiology (168 patients, 22.0%) or surgery/orthopaedics (154 patients, 20.1%). The median Katz ADL score was 0 (IQR 0-2) and 172 patients (25.1%) had cognitive impairment.

Table 1: Baseline characteristics of acutely hospitalized older patients

	N = 765
Age, median (IQR)	79 (74-84)
Male, n (%)	374 (48.9%)
Independent living arrangement, n(%)	698 (91.2%)
High education, n (%)	143 (18.7%)
Academic hospital, n(%)	331 (40.7%)
Clinical specialism, n(%)	
Internal medicine	242 (36.1%)
Cardiology	168 (22.0%)
Surgery/Orthopaedics	154 (20.1%)
Neurology	87 (11.4%)
Pulmonology	73 (9.5%)
Other[1]	41 (5.4%)
Number of medications, median (IQR)	6 (3-8)
History of dementia, n(%)	33 (4.3%)
Katz ADL score, median (IQR)[2]	0 (0-2)
Cognitive impairment, n(%)[3]	172 (25.1%)

IQR = interquartile range, ADL = activities of daily living

[1]Others include gastroenterology, urology, ear nose throat and oncology, all contributing <3.0%.

[2]The Katz ADL score indicates functional status two weeks prior to admission with scores ranging from 0-6. A higher score corresponds with more dependency. In total 15 Katz ADL scores were missing.

[3]Cognitive impairment indicates patients with a 6CIT (six-item cognitive impairment test) score of ≥11. In total 81 6CIT scores were missing.

Figure 1 displays the distribution of ISAR-HP scores. The median ISAR-HP score was 2 (IQR 0-4) and 432 patients (56.5%) were at high risk for functional decline or mortality when using the conventional cut-off of ≥ 2. In total 276 patients (36.1%) experienced functional decline or mortality within 90 days of follow-up. We first performed an external validation of the ISAR-HP. Figure 2 shows the calibration plot of the ISAR-HP for functional decline or mortality. Calibration was insufficient with a Hosmer and Lemeshow goodness-of-fit p-value of 0.007. Predicted probabilities were lower than the observed probabilities with a calibration slope of 0.877 and an intercept of 0.246, indicating an underestimation of the outcome.

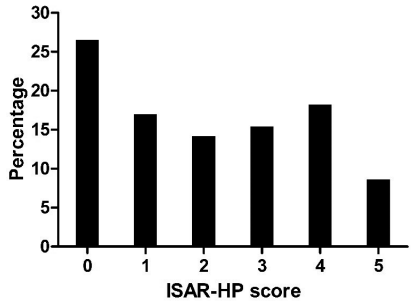

Figure 1: Histogram of ISAR-HP scores

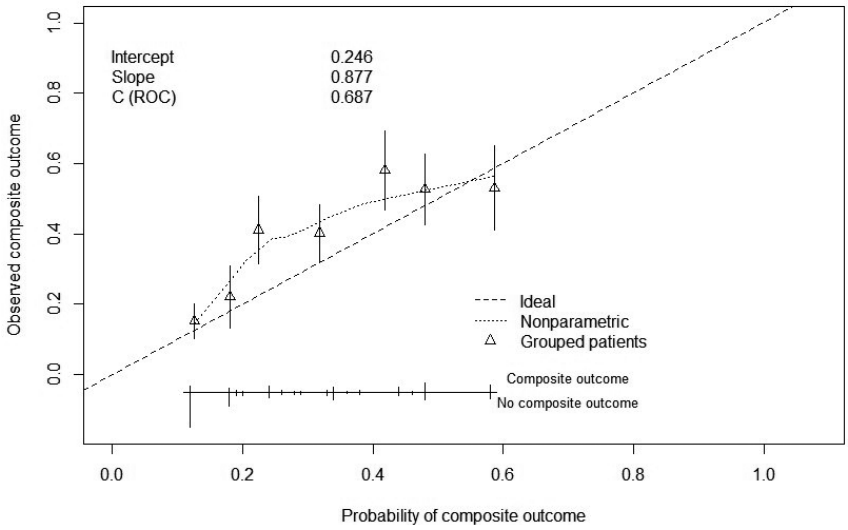

Figure 2: Calibration plot of the predicted probabilities for 90-day functional decline or mortality with the original internal validated ISAR-HP regression equation. The vertical lines represent the relative frequency distribution of predicted probabilities of patients experiencing the combined outcome (above the horizontal line) or not experiencing the outcome (below horizontal line). The triangles represent the grouped patients with 95% confidence intervals. The Hosmer and Lemeshow goodness-of-fit p-value is 0.007

The predictive performance of the ISAR-HP is shown in table 2. Accuracy of the ISAR-HP was reasonable, with an AUC of 0.69 (95%CI 0.65-0.73). External validation of the ISAR-HP for the conventional cut-off of ≥ 2 resulted in a sensitivity of 0.77 (95% CI 0.72-0.82), a specificity of 0.55 (95%CI 0.51-0.60), a PPV of 0.49 (95% CI 0.45-0.54) which results in 45% incorrectly classified high-risk patients and 29% incorrectly classified low-risk patients. Additionally, although of limited clinical applicability, we validated the ISAR-HP for solely functional decline to allow comparison of performance compared with the original study (*Appendix: table 1 and figure 2*). After applying the exclusion criteria for functional decline (113 patients died, 30 patients were lost to follow-up, 17 patients were unable to demonstrate functional decline and 13 patients refused), 592 patients were included of whom 162 (27.4%) experienced functional decline. Discrimination was fair (AUC 0.72, 95% CI 0.67-0.76) and calibration satisfactory (Hosmer and Lemeshow goodness-of-fit p-value 0.068). The PPV and NPV were lower for all cut-offs, with a PPV of 0.42 (95%CI 0.37-0.48) and NPV of 0.89 (95%CI 0.84-0.92) for an ISAR-HP score of ≥ 2.

Table 2: Predictive performance of the ISAR-HP with different cut-off points for the composite outcome 90 days after acute hospitalization in older patients

	High risk n, (%)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% C)	LR- (95% CI)
ISAR-HP ≥1	562 (73.5%)	0.89 (0.84-0.92)	0.35 (0.31-0.40)	0.44 (0.39-0.48)	0.85 (0.79-0.89)	1.37 (1.27-1.48)	0.32 (0.23-0.45)
ISAR-HP ≥2	432 (56.5%)	0.77 (0.72-0.82)	0.55 (0.51-0.60)	0.49 (0.45-0.54)	0.81 (0.76-0.85)	1.72 (1.53-1.94)	0.41 (0.33-0.52)
ISAR-HP ≥3	323 (42.2%)	0.59 (0.53-0.65)	0.67 (0.63-0.71)	0.50 (0.45-0.56)	0.74 (0.70-0.78)	1.80 (1.53-2.12)	0.61 (0.53-0.70)
ISAR-HP ≥4	205 (26.8%)	0.41 (0.35-0.47)	0.81 (0.77-0.84)	0.55 (0.48-0.62)	0.71 (0.67-0.74)	2.13 (1.69-2.69)	0.73 (0.66-0.81)

CI = confidence interval, PPV = positive predictive value, NPV = negative predictive value, LR+ = positive likelihood ratio, LR- = negative likelihood ratio
The area under receiver operating curve of the ISAR-HP model was 0.69 (95% CI 0.65-0.73)

Discussion

The main finding of the present study is that the ISAR-HP with the conventional cut-off of ≥ 2 resulted in more than half of all acutely admitted patients to be considered at high risk for 90-day functional decline or mortality; of these patients 51% did not experience this outcome. Raising the ISAR-HP cut-off to ≥ 4 resulted in a quarter of all patients being classified at high risk and the predictive performance increased marginally.

The ISAR-HP was originally developed to predict 90-day functional decline in patients aged 65 years and older, who were acutely hospitalised for at least 48 hours on a general internal medicine ward.[60] Age and gender were comparable, but compared with our cohort more patients were living dependently (24% in development cohort vs. 9% in APOP cohort) and more patients were cognitively impaired (34% in development cohort vs. 25% in APOP cohort). In the original development study 70% of all older patients were identified as high risk with a PPV of 0.43, implying that 57% of these high-risk patients did not suffer from functional decline. In the APOP cohort 57% were assigned as high risk with a PPV of 0.49, which means 51% showed no decline. The ISAR-HP was also validated for functional decline in cardiac surgery patients aged 65 years and older.[61]Compared with the APOP cohort the patients were younger (mean age 73 years), more often male (64%) and had a better cognitive performance (14% with memory problems). In total 16% of the cardiac surgery patients suffered a functional decline and the AUC of the ISAR-HP was 0.72. The reported PPV of an ISAR-HP score of ≥ 2 was 0.36, which implies that approximately two out of three patients were incorrectly considered high risk. Results from a meta-analysis of screening instruments to predict functional decline and mortality in older patients visiting the emergency department were in line with the results of the ISAR-HP.[4] High-risk groups could not accurately be distinguished from low-risk groups and a relatively high number of patients were incorrectly classified as high risk. In the Netherlands older hospitalised patients are systematically screened for undernutrition, ADL limitations, falls and delirium to prevent 90-day functional decline or mortality: the Safety Management System (VMS+) screening.[65] A third of all VMS+ screened older patients were considered to be at high risk and predictive performance was comparable with our results, with a PPV ranging between 0.50 and 0.57. In summary, the results of our study were in line with ISAR-HP in older patients on both a general internal medicine ward and undergoing cardiac surgery and in line with the performance of the VMS+ screening.

The ISAR-HP was developed and validated to predict only functional decline. In order to replicate that analysis, we had to exclude almost a quarter of patients due to the exclusion criteria, including those who had died during follow-up. As a consequence, this selection no longer reflects clinical practice, where screening is implemented for all patients. A composite outcome of functional decline with mortality as ultimate decline in physical functioning was therefore used. In the present study the predictive performance of different cut-off points for predicting the composite outcome was studied in order to improve efficiency of ISAR-HP screening. Increasing the cut-off point to ≥ 4 resulted in selection of 26.8% of the patients at highest risk and the PPV improved to 55%. Although one in two patients would be inappropriately assigned to the 'high-risk' group, less patients are considered at high risk. To date, the ISAR-HP has been used twice in study settings as a screening instrument for CGA interventions to specifically prevent functional decline in older hospitalised patients. In the Prevention and Reactivation Care Program (PReCaP)

older patients at risk of functional decline received supplementary multidisciplinary, goal-oriented care.[66] n the Transitional Care Bridge Randomised Controlled Trial older patients at risk of functional decline received a systematic CGA, followed by a hospital visit of the community care registered nurse and subsequent multiple home visits after discharge.[67] Although the interventions should be appropriate to prevent functional decline, both the PReCap and the Transitional Care Bridge Randomised Controlled Trial showed no effect on ADL functioning. Based on the findings of our study this may be explained by the fact that many patients were incorrectly selected for the intervention. Taken together, the predictive performance of the ISAR-HP is characterised by low PPVs. Using the ISAR-HP for identifying patients at high risk for functional decline could result in providing inefficient follow-up care.

In screening instruments there is a certain clinical threshold which is determined by the relative weight of false-negative versus false-positive errors. A conservative low cut-off is useful if missing a patient who will undergo a functional decline is more important than incorrect classification of a patient who will not.[38] From the patient perspective a low cut-off is desirable, especially if it results in an intervention without side effects. However, using a stricter cut-off is more useful from another perspective. As an example, in 2012 in total 734,000 patients aged 65 years and over were admitted to hospital in the Netherlands (25% of all older adults)[68] of whom 415,000 might have been at high risk according to our results with the conventional ISAR-HP cut-off of ≥ 2. If the recommended CGA had been performed in all patients, 212,000 CGAs would be carried out unnecessarily in order to prevent the occurrence of the outcome. Using a stricter ISAR-HP cut-off ≥ 4 would result in 197,000 high-risk patients, with 88,000 incorrectly classified as high risk. In patient groups with a lower incidence of functional decline or mortality, such as older cardiac surgery patients, even more interventions will be performed unnecessarily.[61]

Taking into account the limited efficacy and capacity to perform CGAs we therefore recommend to target interventions in a larger group of patients which are inexpensive, less time consuming and not a burden for the patient. A more specific screening instrument for hospitalised older patients is needed to be able to target resource intensive interventions in a smaller group of patients. In our publication on the development and validation of a new screening instrument for older patients visiting the emergency department, we were able to increase specificity compared with the widely accepted ISAR screening instrument. [62] During hospital stay more patient data will become available, such as vital parameters and laboratory results, which may improve predictive performance. Therefore, we are currently developing a new dynamic predicting model for hospitalised older patients.

Several limitations need to be addressed. First, the performance of the ISAR-HP in non-acutely hospitalised older patients, such as elective admissions and via the outpatient

clinic, has not been evaluated. Second, the follow-up data on functional decline were incomplete for 43 patients. From municipal records we know that these patients were alive, which might therefore have resulted in an underestimation of functional decline. The major strength of the study is the inclusion of a representative cohort of older patients. In total 83% of the eligible older patients from different specialisms were included, from both an academic and regional hospital. A second strength is that ISAR-HP was evaluated for the composite outcomes, which reflects predictive performance in clinical practice.

In conclusion, the ISAR-HP with the conventional cut-off of ≥2 incorrectly identifies a large group of patients as being at high risk for functional decline or mortality, and raising the cut-off to 4 only marginally improved performance. Caution is warranted to ensure efficient screening and follow-up interventions.

References

1. IGZ. Basisset Medisch Specialistische Zorg Kwaliteitsindicatoren. 2017 24-02-2017 [cited 2017 24-02]; Available from: *https://www.igz.nl/Images/20160726%20IGZ-Basisset%20MSZ%20 2017%20digitale%20versie%20(enkel)%20v2_tcm294-377263.pdf.*
2. Winograd, C.H., *Targeting strategies: an overview of criteria and outcomes.* J Am Geriatr Soc, 1991. 39 (9 Pt 2): p. 25S-35S.
3. Hoogerduijn, J.G., et al., *The prediction of functional decline in older hospitalised patients.* Age Ageing, 2012. 41 (3): p. 381-7.
4. Hoogerduijn, J.G., et al., *Predicting functional decline in older patients undergoing cardiac surgery.* Age Ageing, 2014. 43 (2): p. 218-21.
5. de Gelder, J., et al., *Predicting adverse health outcomes in older emergency department patients: the APOP study.* Neth J Med, 2016. 74 (8): p. 342-352.
6. Katzman R, B.T., Fuld P, Peck A, Schechter R, Schimmel H., *Validation of a short Orientation-Memory-Concentration Test of cognitive impairment.* Am J Psychiatry., 1983. Jun 140 (6): p. 734-9.
7. Katz, S., et al., *Studies of Illness in the Aged. The Index of Adl: A Standardized Measure of Biological and Psychosocial Function.* JAMA, 1963. (185): p. 914-9.
8. Vergouwe, Y., et al., *Substantial effective sample sizes were required for external validation studies of predictive logistic regression models.* J Clin Epidemiol, 2005. 58 (5): p. 475-83.
9. Hosmer DW , L.S., *Applied logistic regression*. 2nd ed. 2000, New York: Wiley.
10. Team, R.C., *R: A Language and Environment for Statistical Computing*, in *R Foundation for Statistical Computing*. 2016: Vienna, Austria.
11. Carpenter, C.R., et al., *Risk factors and screening instruments to predict adverse outcomes for undifferentiated older emergency department patients: a systematic review and meta-analysis.* Acad Emerg Med, 2015. 22 (1): p. 1-21.
12. Heim, N., et al., *Optimal screening for increased risk for adverse outcomes in hospitalised older adults.* Age Ageing, 2015. 44 (2): p. 239-44.
13. Asmus-Szepesi, K.J., et al., *Evaluation of the Prevention and Reactivation Care Program (PreCaP) for the hospitalized elderly: a prospective nonrandomized controlled trial.* Clin Interv Aging, 2015. (10): p. 649-61.
14. Buurman, B.M., et al., *Comprehensive Geriatric Assessment and Transitional Care in Acutely Hospitalized Patients: The Transitional Care Bridge Randomized Clinical Trial.* JAMA Intern Med, 2016. 176 (3): p. 302-9.
15. Steyerberg, E.W., *Clinical prediction models: a practical approach to development, validation, and updating.*, ed. M.G.K.K.J.S.A.T. W.Wong. 2009, New York: Springer.
16. *Ziekenhuisopnamen; kerncijfers; geslacht, leeftijd en regio*, C.B.v.d. Statitiek, Editor. 2017.

CHAPTER 5

Refinement of the APOP screener

Optimization of the APOP screener to predict functional decline or mortality in older emergency department patients: cross-validation in four prospective cohorts

J. de Gelder, J.A. Lucke, L.C. Blomaard, A.M. Booijen, A.J. Fogteloo, S. Anten, E.W. Steyerberg, J. Alsma, S.C.E. Schuit, A. Brink, B.de Groot, G.J. Blauw, S.P. Mooijaart

Experimental Gerontology 2018; Sep;110:253-259

Abstract

INTRODUCTION: Many screening instruments to predict adverse health outcomes in older patients visiting the emergency department (ED) have been developed, but successful implementation has been hampered because they are insufficiently validated or not tailored for the intended use of everyday clinical practice. The present study aims to refine and validate an existing screening instrument (the APOP screener) to predict 90-day functional decline or mortality in older ED patients.

METHODS: Consecutive older patients (>/=70years) visiting the EDs of four hospitals were included and prospectively followed. First, an expert panel used predefined criteria to decide which independent predictors (including demographics, illness severity and geriatric parameters) were suitable for refinement of the model predicting functional decline or mortality after 90days. Second, the model was cross-validated in all four hospitals and predictive performance was assessed. Additionally, a pilot study among triage nurses experiences and clinical usability of the APOP screener was conducted.

RESULTS: In total 2629 older patients were included, with a median age of 79years (IQR 74-84). After 90days 805 patients (30.6%) experienced functional decline or mortality. The refined prediction model included age, gender, way of arrival, need of regular help, need help in bathing/showering, hospitalization the prior six months and impaired cognition. Calibration was good and cross-validation was successful with a pooled area under the curve of 0.71 (0.69-0.73). In the top 20% patients predicted to be at highest risk in total 58% (95%CI 54%-62%) experienced functional decline or mortality. Triage nurses found the screener well suited for clinical use, with room for improvement.

CONCLUSION: In conclusion, optimization of the APOP screener resulted in a short and more simplified screener, which adequately identifies older ED patients at highest risk for functional decline or mortality. The findings of the pilot study were promising for clinical use.

Introduction

Up to 45% of all older patients experience functional decline or mortality within three months after an emergency department (ED) visit.[1] Multiple screening instruments have been developed to identify older patients at high risk for adverse functional outcomes.[2, 3, 69, 70] Although guidelines include the policy to screen all older patients who visit the ED,[71] these instruments have been rarely implemented as part of routine care. The frequent rejection of developed screening instruments is likely due to poor external validation or the impossibility to integrate the instrument in daily routine care.[72]

The Stiell criteria lists six major methodologic stages to disseminate and implement a developed screener in daily practice.[73] Previously, we have developed and validated the APOP screener to identify patients at risk of mortality or functional decline.[62] The APOP screener is characterised with a high specificity and high positive predictive value, to rule-in patients at highest risk, compared to the well-known Identification of Seniors at Risk (ISAR)[62] and other screening instruments.[4] To improve the chance of successful implementation in clinical practice, a refinement process of the screener is advised.[73] In this process accuracy can be improved, the screener can be simplified and acceptance among the people who have to use the screener in daily practice can be evaluated.

In the present study we aimed to optimize the APOP screener for predicting 90-day functional decline or mortality in older ED patients by selecting predictors based on pre-defined criteria, cross-validation in patients of four hospitals. Additionally, facilitators and barriers of adoption by triage nurses were evaluated in a pilot study.

Methods

Study design and setting

We conducted a multicentre cohort study among consecutive older patients visiting emergency departments (EDs) of four hospitals in the Netherlands: the APOP study[62]. In short, patients were included from September 2014 – November 2014 in the Leiden University Medical Center (LUMC, , Leiden), from March 2015 – June 2015 in Alrijne hospital (Alrijne, Leiderdorp), from May 2016 – July 2016 in Haaglanden Medical Center, location Bronovo (Bronovo, The Hague) and from July 2016 – January 2017 in Erasmus University Medical Center (Erasmus MC, Rotterdam). Training sessions were organized to guarantee that in all hospitals inclusion procedures were equal. During twelve weeks patients were included in the LUMC (7 days a week, 24 hours a day) and in Alrijne hospital (7 days a week, from 10AM-10PM). In Bronovo and Erasmuc MC we aimed to include 500 patients. In Bronovo inclusion was performed 6 days a week, from 10 AM-10 PM and in Erasmus MC 4 days a week (including weekend days) from 10 AM- 10 PM. All patients aged 70 years and over were eligible for inclusion. Exclusion criteria were: red triage category (highest acuity) according to the Manchester Triage System (MTS),[44] an unstable medical condition, no permission of nurse or physician to approach the patient, a language barrier and impossibility to obtain informed consent. The medical ethics committees waived the necessity for formal approval of the study protocol, as the study closely followed routine care. Written informed consent was obtained of all patients or relatives before inclusion.

Baseline

At baseline, data on three domains were assessed. First, demographics including age, sex, living arrangement and level of education. Living arrangement was defined as patients living independent with others, independent alone or in a residential care centre or nursing home. High education includes patients with vocational training or university. Second, severity of disease indicators, included arrival by ambulance, fall related ED visits, triage urgency and chief complains as obtained with the Manchester Triage System (MTS), was scored.[44] The 52 possible chief complaints were classified into seven main groups (supplementary table 1). Third, the geriatric parameters included the presence of polypharmacy, use of walking device, Katz ADL (activities of daily living) score[46] and cognition measured by the six item Cognitive Impairment Test (6-CIT).[45] Polypharmacy was defined as the use of five or more different medications at home, self-reported by the patient. The Katz ADL evaluates the ADL situation two weeks prior to the ED visit with six yes/no questions on basic activities of daily living (zero to six point scale). Higher scores indicate more dependency. The 6-CIT is a short cognition test with scores ranging from 0-28, with a score of 11 or higher indicating moderate to severe cognitive impairment, comparable to an MMSE of 24 or lower.[48] To reduce the number of questions needed to be asked to test cognition, all possible single 6-CIT questions were combined. The

combination with the highest Phi correlation coefficient was selected (supplementary table 2). Patients were considered cognitively impaired if they incorrectly answered the question 'what year is it now?' and/or 'say the months in reverse order' (incorrect if two or more errors in months). If the patient is diagnosed with dementia or if it is impossible to obtain answers for the two questions for any reason (e.g. due to mental status), cognition was also considered to be impaired.

Outcome

The primary adverse health outcome was the composite outcome of functional decline or mortality at 90 days follow-up, equal to the development study and to ensure that screening can be implemented for all geriatric ED patients.[62, 74] Mortality was incorporated into the composite outcome, as it can be seen as ultimate decline. Functional decline was defined as at least one point increase in Katz ADL score or new institutionalization (e.g. nursing home admission) at 90 days after ED visit. To obtain follow-up data, patients were contacted by telephone 90 days after the ED index visit. In case of no response after three attempts the general practitioner was contacted to verify phone number and living arrangement (new institutionalization). Finally, to patients who could not be contacted, a letter was sent with a request for a written response. Data on mortality were obtained from the municipal records.

1. Refinement of predictors in the model

The original APOP screener, which predicts 90-day functional decline or mortality and solely 90-day mortality, was developed with data of LUMC patients and validated with Alrijne patients.[62]. For refining of the model, instead of redeveloping the APOP screener with regression techniques, selection criteria were formulated to select predictors[Box 1].[38] The five most important criteria for a strong and user friendly prediction model were extracted and reformatted with permission. Consensus to meet all five criteria of predictors was obtained in a multidisciplinary meeting consisting of physicians (emergency medicine, internal medicine and geriatrics), nurses (emergency medicine, internal medicine and geriatrics) and a statistician.

Box 1: Selection criteria for predictors

Selection criteria	Explanation
Applicable	The collection and definition of predictors should follow routine clinical care as good as possible and require as little extra work as possible
Reliably measured	Objective and robust predictor to reduce inter-observer variability or variability between different hospitals.
Easily measured	Predictor should be fast and easy to obtain, to ensure screening can be finished in short time.
Early available	Predictor should be available at the moment of triage of the patient
Strong predictors	Based on the strength of association with outcome Based on the prevalence of predictor. A wide distribution is preferred over a narrow distribution.

2. Cross-validation of the screener

The final selection of predictors represent the APOP screener and were cross-validated in four hospitals. The LUMC is an academic hospital in with a level 1 trauma centre and Alrijne hospital is a community hospital with a level 2 trauma centre. Both hospitals are located in a small city. The Bronovo hospital is an community hospital with a level 2 trauma centre. The Bronovo hospital is located in a district with relatively many wealthy older people. In the region patients with a suspicion of hip fracture will be sent to the Bronovo. The Erasmus MC is an academic hospital with a level 1 trauma centre and located in the centre of a big city.

3. Pilot study for usability and acceptance of the screener

Eight triage nurses were instructed to use the refined APOP screener for one week in patients aged 70 and over to track the time needed to complete the screening and evaluate usability. Afterwards, an evaluation form was sent to the nurses to get an first impression of possible barriers and clinical application of the APOP screener. A five-level Likert scale was used to score results with the possibility to score strongly disagree (1), disagree (2), neither agree nor disagree (3), agree (4) and strongly agree (5). It was possible to write down additional feedback in free text.

Statistical analysis

Baseline descriptive characteristics are presented as numbers with percentages (%) and median with inter quartile ranges (IQR). Multivariable binary logistic regression was used to estimate the regression coefficients of the prediction model for 90-day functional decline or mortality. Calibration of the prediction model was graphically displayed with calibration plots.[75] A minimum number of 10 events per candidate predictor was used to obtain good predictions with adequate statistical power.[38] Validity of the model was assessed with an internal-external validation design.[76] The robustness of the model was evaluated by a leave-one-hospital-out cross-validation procedure, with patients of each single hospital representing the validation cohort for a model based on the patients of the other three hospitals.[77] External validity was assessed by pooling the cross-validated area under the receiver operating characteristic curves (AUCs) of the four single hospitals using a random-effect meta-analysis.[77] Predictive performance of the model was evaluated for the patients with the highest 30%, 20% and 10% predicted risk, with sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+) and negative likelihood ratio (LR-). Additionally, the prediction model, calibration plot and predictive performance for solely 90-day mortality was assessed. Mean Likert scale scores with standard deviation (SD) were used to analyse usability of the screener. Analysis was performed with IBM SPSS statistics version 23 and R software (version 3.1.1.)

Results

A total of 3,544 individual patients aged 70 years and older visited the emergency departments (EDs) of the four hospitals combined during the inclusion of the study period. Of those, 3,147 were eligible for inclusion in the APOP study. In total 2,629 patients were included (84% of the eligible patients (figure 1)).

Table 1 shows the baseline characteristics of the APOP study population and stratified per study centre. The median age was 79 years (IQR 74-84) for the combined group, ranging from a median of 76 years in the Erasmus MC to median 82 years in Bronovo. In total 1,236 patients (47.0%) were male, 1,339 patients (50.9%) arrived by ambulance and 659 patients (25.1%) experienced a fall prior to the ED visit. Polypharmacy was found in 1,552 patients (57.9%). Impaired cognition was present in 492 patients (20.5%).

Table 1: Baseline characteristics of older patients visiting the emergency department

	All (N = 2629)	LUMC (N = 751)	Alrijne (N = 881)	Bronovo (N = 498)	Erasmus MC (N = 499)
Demographics					
Age (years), median (IQR)	79 (74-84)	78 (74-83)	80 (75-84)	82 (75-87)	76 (73-80)
Male, n(%)	1236 (47.0%)	362 (48.2%)	427 (48.5%)	164 (32.9%)	283 (56.7%)
Living arrangement, n (%)					
Independent with others	1421 (54.1%)	414 (55.1%)	498 (56.5%)	208 (41.8%)	301 (60.4%)
Independent alone	991 (37.7%)	274 (36.5%)	314 (35.6%)	231 (46.4%)	172 (34.5%)
Residential care or nursing home	216 (8.2%)	63 (8.4%)	69 (7.8%)	59 (11.8%)	25 (5.0%)
High educated, n (%)	586(22.4%)	155 (20.6%)	164 (18.6%)	147 (29.6%)	120 (24.3%)
Severity of disease indicators					
Arrival by ambulance, n (%)	1339 (50.9%)	405 (53.9%)	432 (49.0%)	256 (51.4%)	246 (49.3%)
Triage urgency, n (%)					
> 1 hour (green)	717 (27.3%)	159 (21.2%)	353 (40.1%)	104 (20.9%)	101 (20.2%)
< 1 hour (yellow)	1534 (58.3%)	391 (52.1%)	470 (53.3%)	347 (69.7%)	326 (65.3%)
< 10 min (orange)	378 (14.4%)	201 (26.8%)	58 (6.6%)	47 (9.4%)	72 (14.4%)
Chief complaint, n (%)					
Minor trauma	815 (31.0%)	218 (29.0%)	232 (26.3%)	232 (46.6%)	133 (26.7%)
Malaise	465 (17.7%)	137 (18.2%)	176 (20.0%)	85 (17.1%)	67 (13.4%)
Chest pain	393 (14.9%)	111 (14.8%)	167 (19.0%)	57 (11.4%)	58 (11.6%)
Dyspnea	320 (12.2%)	76 (10.1%)	131 (14.9%)	43 (8.6%)	70 (14.0%)
Abdominal pain	282 (10.7%)	84 (11.2%)	96 (10.9%)	35 (7.0%)	67 (13.4%)
Loss of consciousness	146 (5.6%)	49 (6.5%)	38 (4.3%)	14 (2.8%)	45 (9.0%)
Others	208 (7.9%)	76 (10.1%)	41 (4.7%)	32 (6.4%)	59 (11.8%)
Fall prior to ED visit, n (%)	659 (25.1%)	211 (28.1%)	192 (21.8%)	179 (35.9%)	77 (15.4%)
Geriatric measurements					
Polypharmacy, n (%)	1552 (57.9%)	441 (58.7%)	509 (57.8%)	241 (48.4%)	331 (66.3%)
Use of walking device, n (%)	1114 (42.5%)	302 (40.2%)	378 (42.9%)	243 (48.9%)	191 (38.4%)
Katz ADL score, median (IQR)	0 (0-1)	0 (0-1)	0 (0-1)	0 (0-2)	0 (0-1)
Impaired cognition, n (%)	492 (20.5%)	140 (19.9%)	174 (21.6%)	111 (23.9%)	67 (15.9%)

N = number, IQR= Interquartile range, ADL = activities of daily living, ED = emergency department
Missings
LUMC: 5 level of education, 4 walking device, 6 Katz ADL score, 47 impaired cognition
Alrijne: 3 level of education, 3 walking device, 22 Katz ADL score, 75 impaired cognition
Bronovo: 2 level of education, 1 walking device, 3 Katz ADL score, 33 impaired cognition
Erasmus: 1 living arrangement, 6 level of education, 2 walking device, 9 Katz ADL score, 77 impaired cognition

1. Refinement of predictors in the model

Table 2 shows the results of the selection of the predictors based on the predefined criteria. The APOP screener consists of seven predictors which meet all criteria: age, gender, arrival by ambulance, need of regular help (IADL), need for help with bathing or showering, hospitalization in the prior 6 months and impaired cognition. Arguments of ineligibility of the other predictors can be found in supplementary table 3.

Table 2: Selection of predictors for refinement of the APOP screener

	Applicable	Reliably measured	Easily measured	Readily available	Strong predictor
Age	+	+	+	+	+
Gender	+	+	+	+	+
Living arrangement	-	+	+	+	+
Level of education	+	+	+	+	-
Arrival by ambulance	+	+	+	+	+
Triage category	+	-	+	+	-
Chief complaint	+	-	+	+	+
Fall prior to ED visit	+	-	-	+	+
Vital measurements	+	+	+	-	-
Laboratory results	+	+	+	-	+
Polypharmacy	+	-	-	+	+
Use of walking device	+	+	+	+	-
Need regular help (IADL)	+	+	+	+	+
Need help bathing showering	+	+	+	+	+
Need help dressing	+	+	+	+	-
Hospitalized past 6 months	+	+	+	+	+
Impaired cognition	+	+	+	+	+

In bold: eligible predictors

ED: Emergency department, IADL: Instrumental activities of daily living

Table 3: Prediction model for 90-day functional decline or mortality in older patients visiting the emergency department

	OR (95% CI)
Age (per 5 years increase)	1.30 (1.21-1.40)
Male	0.93 (0.78-1.12)
Arrival by ambulance	1.58 (1.32-1.91)
Need help prior to ED visit (IADL)	1.71 (1.39-2.10)
Need help bathing or showering	1.76 (1.40-2.21)
Hospitalized past six months	1.54 (1.27-1.87)
Impaired cognition	1.29 (1.06-1.57)

OR = odds ratio, ED = emergency department, IADL = instrumental activities of daily living
Equation:
1/(1+exp(-(-5.848 + 0.262 x 'age/5)' + -0.072 x 'male' + 0.460 x 'arrival by ambulance' + 0.534 x 'need help prior to ED visit' + 0.567 x 'need help bathing or showering' + 0.432 x 'hospitalized past six months' + 0.255 x 'impaired cognition')))
Application: http://screener.apop.eu

2. Cross-validation of the screener

A total of 139 out of 2,629 patients (5.3%) were lost to follow-up for data on physical functioning, but from municipal records we verified that they were alive. The incidence of 90-day composite outcome in the study population was 30.6% (805 out of 2,629 patients) (supplementary figure 1). Table 3 shows the result of the multivariable logistic regression of the refined screener. All selected predictors, except gender, were statistically significant associated with the outcome. The individual predicted risk of a patient to experience the outcome can be calculated by using the equation in the legend of the table or by using a free web-based calculator: http://screener.apop.eu/. Cross-validation of the screener was successful, with comparable AUC's between the four individual hospitals (figure 2). External validity of the screener was good, with a pooled AUC of 0.71 (95%CI 0.69-0.73). The predicted probabilities were in line with the observed, as can be seen in the calibration plot (supplementary figure 2). Predictive performance for 90-day functional decline or mortality is shown for the 30%, 20% and 10% patients at highest risk (table 4). Stricter thresholds for high risk increased specificity, positive predictive value (PPV) and positive likelihood ratio (LR+). The PPV for 90-day functional decline or mortality was 0.53 (95%CI 0.49-0.56) in the 30% patients at highest risk, 0.58 (95%CI 0.54-0.62) in the 20% patients at highest risk and 0.60 (95%CI 0.54-0.66) in the 10% patients at highest risk.

An additional analysis was performed to predict 90-day mortality as a separate end point (supplementary material). In total 9.9% of the patients (259 out of 2,629) deceased within 90 days after visiting the emergency department (supplementary figure 1). Accuracy of the refined screener was good with an AUC of 0.74 (95%CI 0.71-0.77) (supplementary table 3), calibration was successful (supplementary figure 2) and the PPV ranged from 0.20 (95%CI

0.17-0.23) for the 30% patients at highest risk to 0.28 (95%CI 0.23-0.34) for the 10% patients at highest risk (supplementary table 4).

3. Usability and acceptance of the screener in the pilot study

A total of 60 patients was screened by eight triage nurses. The mean time to complete the screener was 93 seconds (SD 29). The overall rating of clinical usability was positive, with a mean Likert score of 3.79 (SD 0.63) (supplementary table 5). The screener was easy to administer, the triage nurses found it important to screen and experienced no big burden for the patient. In current form some nurses experienced an increase in workload. These nurses advised that workload can be reduced by incorporating the APOP-screener in the electronic patient files instead of using the web-based application.

Table 4: Predictive performance of final prediction model for 90-day functional decline or mortality (N=2608)

	Number of patients at risk	Sens (95% CI)	Spec (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% CI)	LR- (95% CI)
30% at highest risk	780	0.52 (0.48-0.55)	0.80 (0.78-0.81)	0.53 (0.79-0.56)	0.79 (0.77-0.81)	2.51 (2.24-2.81)	0.61 (0.57-0.66)
20% at highest risk	521	0.38 (0.35-0.41)	0.88 (0.86-0.89)	0.58 (0.54-0.62)	0.76 (0.74-0.78)	3.15 (2.71-3.67)	0.71 (0.67-0.74)
10% at highest risk	260	0.20 (0.17-0.23)	0.94 (0.93-0.95)	0.60 (0.54-0.66)	0.73 (0.71-0.74)	3.40 (2.69-4.30)	0.85 (0.82-0.88)

Sens = sensitivity, Spec = specificity, PPV = positive predictive value, NPV = negative predictive value, LR+ = positive likelihood ratio, LR- = negative likelihood ratio, CI = confidence interval

Discussion

The screener was refined by selecting predictors based on predefined criteria for predicting 90-day functional decline or mortality in older emergency department patients. The refined model was cross-validated in four hospitals and showed satisfactory discrimination and calibration. Predictive performance was good, with high positive predictive values. A pilot performed by triage nurses showed adequate usability of the screener in clinical practice, with room for improvement.

In the present study the screener was refined in order to increase its usefulness in clinical practice. In a multidisciplinary meeting predictors were chosen with predefined generally accepted criteria,[38] which took into account both the association with the outcome and possible barriers for implementation. Compared to the original model, gender and cognition were added and number of medications was removed. Gender is readily information upon attendance and associated with the outcome.[62, 78] Impaired

cognition is highly prevalent in the ED,[18, 20] and frequently underdiagnosed[79] and is associated with functional decline.[80, 81] Although number of medications is known to be associated with functional decline and mortality,[62] the predictor was not selected for other reasons. Inter observer variability can easily be introduced due to the combined medications of different pharmacological sub classifications or prescribed 'as-needed' and patients tend to hand over pill boxes, which takes too much time at the moment of triage. At the end, the refinement process resulted in a more simplified screener, based on a large heterogenetic group of older patients.

The refined APOP screener was successful cross-validated in four different hospitals, with universal predictors . The study population was representative, with a high proportion of included patients. We therefore assume that the screener is generalizable for EDs in Western countries, but needs to be external validated for confirmation first. Predictive performance of the APOP screener differs compared to the Identification of Seniors at Risk (ISAR) tool[2] and Triage Risk Screening Tool (TRST)[3]. Sensitivity of the ISAR and TRST are higher (pooled estimate 0.79 and 0.66) and both the specificity (pooled estimate 0.37 and 0.47) and positive likelihood ratio (pooled estimate 1.25 and 1.23) are lower.[4] Although a higher sensitivity will including more patients who will decline, the increased risk to experience the composite outcome for the 'high risk' group by using the screener is minimal. According to these estimates, given the baseline risk of 30% for experiencing the composite outcome, patients with a positive ISAR or TRST screening have a risk of 35% to experience the outcome. We suggest to effectively select patients at highest risk, enabling clinicians to take measures in a smaller group of patients with a higher risk of a potential adverse outcome. The cut-off was therefore set for the 20% patients at highest risk.[56] The risk of experiencing functional decline or mortality in this high risk group increases from 30% (incidence) to 58% (PPV).

Usability of the screener was evaluated among triage nurses in a pilot. With a mean time of 93 seconds to complete screening, the APOP screener is now shorter compared to the original screener. Although the screener was easy to administer and no burden for the patient, suggestions for improvement were given. Some triage nurses experienced difficulties in obtaining the screening result via the web-based application. To make the screener more applicable for routine care, the screener needs to be integrated in the electronic patient files. Second, no follow-up interventions were conducted after screening yet, which ensures that some nurses experienced that workload rather increased than decreased. As an example, in order to reduce the ED length of stay, a fast-track admission trajectory can be developed in high risk patients who need to be hospitalized. We are currently developing a concomitant educational program to train medical personal and will take the feedback into account.

The APOP screener has been prospectively validated[62] and in the present study the screener is successfully refined to increase its usefulness in clinical practice while preserving predictive performance. The next step is to implement the APOP screener in clinical practice. In addition, an implementation study will be conducted to translate the research into clinical practice and to achieve acceptance of the screener of involved stakeholders. At the same moment the educational program will be disseminated to increase awareness of all health care professionals, of which low-risk patients also will benefit. In patients at high risk for functional decline or mortality and in patients with impaired cognition, follow-up actions and interventions will be conducted[Box 2]. Effect of the interventions on outcomes, including quality of life, will be evaluated.[82] After patient and physicians acceptance is evaluated, the balance between 'costs' and 'benefits'[41] will be investigated and a strategy for wide-spread dissemination and implementation will be developed.

Box 2: Overview of possible actions and interventions after screening result

	High risk functional decline or mortality	**Impaired cognition**
Emergency Department		
(triage) Nurse	Informs involved health care professionals If patient is alone, ask family member or care giver to come to the ED. Nurses patient on a comfortable bed	Informs involved health care professionals If patient is alone, ask family member or care giver to come to the ED. Nurses patient on a comfortable bed Starts multicomponent delirium prevention measures
(ED) Physician	Takes the screening result into account in the diagnostic process (e.g. screen for delirium) and decision making.	
Patients discharged home		
(triage) Nurse	Put patient on the list to call back the next day to verify status and to answer questions	Put patient on the list to call back the next day to verify status and to answer questions
(ED) Physician	Informs general practitioner (by telephone or email)	Hands over paper discharge instructions Informs general practitioner (by telephone or email)
Patients admitted to the hospital		
(triage) Nurse	Informs colleague Invites family member or care giver to stay with the patient during transfer	Informs colleague Invites family member or care giver to stay with the patient during transfer
(ED) Physician	Informs colleague Ask geriatric liaison service in consultation	Informs colleague Ask geriatric liaison service in consultation

Some limitations have to be addressed. First, we were not able to investigate all potentially important determinants of the composite outcome (e.g. malnutrition or the presence of care givers). Second, the screener needs further validation to obtain performance in other countries. Third, the pilot study has insufficient power to draw firm conclusions and did not test the effect of applying measures in high-risk patients. Currently we are conducting a large implementation study of the refined APOP screener. Our study has several strengths. First, a large unselected group of older patients visiting the ED of four hospitals was included (84%) with a high follow-up rate (95%). Second, the prospective design of the study enabled to take important geriatric parameters, such as cognition, into account. Third, the differences between baseline characteristics in study centers and the internal-external validation design enabled to use as much possible data to increase generalisability of the screener.

In conclusion, optimization of the APOP screener resulted in a short and more simplified screener, which adequately identifies older ED patients at highest risk for functional decline or mortality. The findings of the pilot study were promising for clinical use.

References

1 Aminzadeh F, Dalziel WB. Older adults in the emergency department: a systematic review of patterns of use, adverse outcomes, and effectiveness of interventions. *AnnEmergMed*. 2002;39: 238-247.

2 McCusker J, Bellavance F, Cardin S, Trepanier S, Verdon J, Ardman O. Detection of older people at increased risk of adverse health outcomes after an emergency visit: the ISAR screening tool. *JAmGeriatrSoc*. 1999;47: 1229-1237.

3 Meldon SW, Mion LC, Palmer RM*, et al.* A brief risk-stratification tool to predict repeat emergency department visits and hospitalizations in older patients discharged from the emergency department. *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine*. 2003;10: 224-232.

4 Rowland K, Maitra AK, Richardson DA, Hudson K, Woodhouse KW. The discharge of elderly patients from an accident and emergency department: functional changes and risk of readmission. *Age Ageing*. 1990;19: 415-418.

5 Runciman P, Currie CT, Nicol M, Green L, McKay V. Discharge of elderly people from an accident and emergency department: evaluation of health visitor follow-up. *Journal of advanced nursing*. 1996;24: 711-718.

6 Carpenter CR. ACEP guidelines.

7 Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. *BMJ*. 2009;338: b606.

8 Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Annals of emergency medicine*. 1999;33: 437-447.

9 de Gelder J, Lucke JA, de Groot B*, et al.* Predicting adverse health outcomes in older emergency department patients: the APOP study. *The Netherlands journal of medicine*. 2016;74: 342-352.

10 Carpenter CR, Shelton E, Fowler S*, et al.* Risk factors and screening instruments to predict adverse outcomes for undifferentiated older emergency department patients: a systematic review and meta-analysis. *Academic emergency medicine : official journal of the Society for Academic Emergency Medicine*. 2015;22: 1-21.

11 Mackway-Jones K. Manchester Triage Group. *Emergency Triage*. 1997.

12 Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of Illness in the Aged. The Index of Adl: A Standardized Measure of Biological and Psychosocial Function. *JAMA : the journal of the American Medical Association*. 1963;185: 914-919.

13 Katzman R BT, Fuld P, Peck A, Schechter R, Schimmel H. Validation of a short Orientation-Memory-Concentration Test of cognitive impairment. *Am J Psychiatry*. 1983;Jun: 734-739.

14 Tuijl JP, Scholte EM, de Craen AJ, van der Mast RC. Screening for cognitive impairment in older general hospital patients: comparison of the Six-Item Cognitive Impairment Test with the Mini-Mental State Examination. *International journal of geriatric psychiatry*. 2012;27: 755-762.

15 de Gelder J, Haenen E, Lucke JA*, et al.* Optimising the ISAR-HP to screen efficiently for functional decline in older patients. *The Netherlands journal of medicine*. 2017;75: 379-385.

16 Steyerberg EW. *Clinical prediction models: a practical approach to development, validation, and updating*. New York: Springer, 2009.
17 Austin PC, Steyerberg EW. Graphical assessment of internal and external calibration of logistic regression models by using loess smoothers. *Statistics in medicine*. 2014;33: 517-535.
18 Steyerberg EW, Harrell FE, Jr. Prediction models need appropriate internal, internal-external, and external validation. *Journal of clinical epidemiology*. 2016;69: 245-247.
19 Austin PC, van Klaveren D, Vergouwe Y, Nieboer D, Lee DS, Steyerberg EW. Geographic and temporal validity of prediction models: different approaches were useful to examine model performance. *Journal of clinical epidemiology*. 2016;79: 76-85.
20 Sirois MJ, Emond M, Ouellet MC*, et al.* Cumulative incidence of functional decline after minor injuries in previously independent older Canadian individuals in the emergency department. *Journal of the American Geriatrics Society*. 2013;61: 1661-1668.
21 Schofield I, Stott DJ, Tolson D, McFadyen A, Monaghan J, Nelson D. Screening for cognitive impairment in older people attending accident and emergency using the 4-item Abbreviated Mental Test. *European journal of emergency medicine : official journal of the European Society for Emergency Medicine*. 2010;17: 340-342.
22 Litovitz GL, Hedberg M, Wise TN, White JD, Mann LS. Recognition of psychological and cognitive impairments in the emergency department. *The American journal of emergency medicine*. 1985;3: 400-402.
23 Hustey FM, Meldon SW. The prevalence and documentation of impaired mental status in elderly emergency department patients. *Annals of emergency medicine*. 2002;39: 248-253.
24 Provencher V, Sirois MJ, Ouellet MC*, et al.* Decline in activities of daily living after a visit to a canadian emergency department for minor injuries in independent older adults: are frail older adults with cognitive impairment at greater risk? *Journal of the American Geriatrics Society*. 2015;63: 860-868.
25 Lucke JA, Gelder J, Clarijs F*, et al.* Early prediction of hospital admission for emergency department patients: a comparison between patients younger or older than 70 years. *Emergency medicine journal : EMJ*. 2017.
26 de Gelder J, Lucke JA, Heim N*, et al.* Predicting mortality in acutely hospitalized older patients: a retrospective cohort study. *Internal and emergency medicine*. 2016;11: 587-594.
27 Van Deudekom FJ, De Gelder J, Lucke JA*, et al.* Determinants of self-rated health in older adults before and 3 months after an emergency department visit: a prospective study. *European journal of emergency medicine : official journal of the European Society for Emergency Medicine*. 2018.
28 Mooijaart SP, Broekhuizen K, Trompet S*, et al.* Evidence-based medicine in older patients: how can we do better? *The Netherlands journal of medicine*. 2015;73: 211-218.

CHAPTER 6

Predictors and outcomes of revisits

Predictors and Outcomes of Revisits in Older Adults Discharged from the Emergency Department

J. de Gelder, J.A. Lucke, B. de Groot, A.J. Fogteloo, S. Anten, C. Heringhaus, O.M. Dekkers, G.J. Blauw, S.P. Mooijaart

Published in Journal of the American Geriatrics Society 2018; 66(4):735-74

Abstract

OBJECTIVES: To study predictors of emergency department (ED) revisits and the association between ED revisits and 90-day functional decline or mortality.

DESIGN: Multicenter cohort study.

SETTING: One academic and two regional Dutch hospitals.

PARTICIPANTS: Older adults discharged from the ED (N=1,093).

MEASUREMENTS: At baseline, data on demographic characteristics, illness severity, and geriatric parameters (cognition, functional capacity) were collected. All participants were prospectively followed for an unplanned revisit within 30 days and for functional decline and mortality 90 days after the initial visit.

RESULTS: The median age was 79 (interquartile range 74-84), and 114 participants (10.4%) had an ED revisit within 30 days of discharge. Age (hazard ratio (HR)=0.96, 95% confidence interval (CI)=0.92-0.99), male sex (HR=1.61, 95% CI=1.05-2.45), polypharmacy (HR=2.06, 95% CI=1.34-3.16), and cognitive impairment (HR=1.71, 95% CI=1.02-2.88) were independent predictors of a 30-day ED revisit. The area under the receiver operating characteristic curve to predict an ED revisit was 0.65 (95% CI=0.60-0.70). In a propensity score-matched analysis, individuals with an ED revisit were at higher risk (odds ratio=1.99 95% CI=1.06-3.71) of functional decline or mortality.

CONCLUSION: Age, male sex, polypharmacy, and cognitive impairment were independent predictors of a 30-day ED revisit, but no useful clinical prediction model could be developed. However, an early ED revisit is a strong new predictor of adverse outcomes in older adults.

Introduction

Emergency department (ED) use by older adults is increasing, and up to 22% have a revisit to the ED within 1 month after discharge from the ED.[83-88] Some stakeholders view early ED revisits unfavorably, and they are sometimes used as a proxy for preventable adverse outcomes. Revisit might be preventable if it were possible to identify a group of individuals with greater risk of returning to the ED at their first visit. Published studies have had inconsistent results in identifying risk factors.[85, 89, 90]

International guidelines to provide optimal emergency care to older adults include recommendations for safe discharge planning.[91, 92] The recommendations positively influence the care of older adults and contribute to better outcomes. In the first 3 months after an ED visit, older adults are at greater risk of functional decline and mortality. [1] It can be speculated that older adults who need to return to the ED early may even be at greater risk of adverse health outcomes, but to our knowledge this has not been studied.

We conducted a cohort study in individuals aged 70 and older visiting 3 EDs in the Netherlands. The aim of the present study was to identify predictors of a 30-day ED revisit and to investigate the association between an early ED revisit and the composite outcome of 90-day functional decline or mortality.

Methods

Study design and setting

We performed a cohort study in older adults visiting EDs in the Netherlands, the Acutely Presenting Older Patients Study.[62] Older adults who presented to the Leiden University Medical Center (LUMC, Leiden) ED from September to November 2014, the Alrijne Hospital (Leiderdorp) from March to June 2015, and the Haaglanden Medical Center (HMC Bronovo, The Hague) from May to July 2016 were included. Inclusion was conducted during a 3-month period in the LUMC (7 days a week, 24 hours a day) and in the Alrijne hospital (7 days a week from 10:00 a.m. to 10:00 p.m.). In the HMC Bronovo, we aimed to include a total of 500 participants, 6 days a week from 10:00 a.m. to 10:00 p.m. The eligibility criteria were not restrictive; all individuals aged 70 and older who visited the ED during the predefined period were included. Exclusion criteria were red triage category according to the Manchester Triage System[44] (highest acuity), unstable medical condition, no permission of nurse or physician to approach the individual, language barrier, and inability to provide informed consent. Relatives were allowed to answer questions, except for the cognition test. For the present study, all individuals discharged from the ED who had not had a prior ED visit within 30 days were used. The medical ethics committees of all hospitals waived the necessity for formal approval of the present study, because the study closely follows routine care. All participants or authorized relatives provided written informed consent before inclusion. The Strengthening the Reporting of Observational Studies in Epidemiology[93] and Transparent Reporting of a Multivariate Prediction Model for Individual Prognosis or Development reporting guidelines were used.[94]

Revisits

An early ED revisit was defined as an unscheduled ED revisit during the 30 days after discharge from the ED that was not part of prearranged care (e.g., ED return on doctor's request, e.g., during the weekend).[95] Records were checked to assess data on revisits. To reduce the risk of missing revisits to other hospitals, after 90 days of follow-up, participants were asked over the telephone whether they had made an ED visit to another hospital after the index visit, and information was subsequently verified. In the Netherlands, individuals who return to the hospital after being discharged are always brought to the ED first. Training was provided before the telephone calls were made to make sure that callers understood the script. JdG, JL, and a medical student collected follow-up information.

Potential predictors of ED revisits

Participants had to answer a limited set of questions within 30 to 45 minutes after arrival at the ED. Data were collected in 3 categories. First, information was collected on demographic characteristics, including age, sex, living arrangement, and level of education. Living arrangement was defined as living independently with others,

independently alone, or in a residential care center or nursing home. Vocational training or university was considered high education. Second, severity of medical condition was assessed according to arrival by ambulance, fall-related ED visit, triage category, and chief complaint determined according to the Manchester Triage System (MTS). [44] The MTS was part of routine clinical care in all hospitals and includes standardized series of flow charts for various presentations to determine urgency in a 5-level triage color, comparable with the Emergency Severity Index.[96] The 52 MTS chief complaints were classified into 7 main groups (Supplementary Table S1). Third, geriatric parameters were polypharmacy, use of a walking device, Katz activity of daily living (ADL) score,[46] and cognition measured using the 6-item Cognitive Impairment Test (6-CIT).[45] Polypharmacy was defined as the use of 5 or more different medications at home, self-reported by the individual. The Katz ADL questionnaire consists of 6 yes-or-no questions on basic ADLs, with higher scores indicating greater dependency. We asked about ADL ability 2 weeks before the ED visit to exclude negative effects of the acute medical complaint on functional status. The 6-CIT is a short cognition test with scores ranging from 0 to 28. A score of 11 or higher is comparable with the conventional cut-off of 23 or lower on the Mini-Mental State Examination and indicates moderate to severe cognitive impairment, irrespective of the cause (e.g., delirium, dementia, depression, medical illness).[48, 97, 98] Two authors (JdG, JL) and trained medical students prospectively collected patient data using structured data collection forms at all three hospitals. Training sessions occurred before the beginning of the study to ensure consistency in questionnaire administration.

Outcomes

The primary adverse health outcome was the composite outcome of functional decline or mortality at 90 days of follow-up Functional decline was defined as a 1-point or greater increase in Katz ADL score or new institutionalization 90 days after the ED visit. Mortality was incorporated into the composite outcome because to omit death would ignore the ultimate functional decline. Data on mortality were obtained from municipal records. Participants were contacted by telephone 90 days after the ED index visit, and a limited set of questions was asked, including living arrangement, functioning, and ED revisits in other hospitals. If there was no response after 3 attempts, the general practitioner was contacted to verify the telephone number and living arrangement (new institutionalization). Finally, a letter was sent to participants who could not be contacted with a request for a written response.

Statistical analysis

Descriptive characteristics are presented as medians with interquartile ranges (IQRs) and numbers with percentages. P-values were calculated using the chi-square test for categorical data and the Mann-Whitney U test for non-normally distributed numeric data. Missing data are reported in the table footnotes. A Kaplan-Meier curve was used to show

the cumulative incidence of 30-day ED revisits. Multivariable Cox regression was used to investigate the association between all baseline characteristics at the index visit and a 30-day ED revisit. The proportional hazards assumption for each prognostic factor was tested using time-dependent covariates, and a prediction model was developed with logistic regression using backward elimination with the Akaike Information Criterion. Ten events per candidate predictor are required to obtain adequate predictions.[38] Discrimination was assessed by calculating the area under the curve (AUC), and calibration was quantified using the Hosmer and Lemeshow test.[52] Predictive performance for the participants with the highest 30%, 20%, and 10% predicted risk was evaluated according to sensitivity, specificity, positive predictive value, negative predictive value, positive likelihood ratio (LR+), and negative likelihood ratio (LR–), with 95% confidence intervals (CIs). A propensity score–matched analysis was conducted to account for the risk of an ED revisit.[99] First a propensity score was estimated for early revisit from the entire population by including the variables demographic characteristics, severity of medical condition, and geriatric parameters. Second, controls were matched at a 1:1 ratio with a propensity score tolerance of 0.05 to minimize potential confounding and selection biases.

The association between an early ED revisit and 90-day adverse health outcome was assessed using binary logistic regression analysis using the group with early ED revisits and the matched control group. Predictive performance of an early revisit was shown using sensitivity, specificity, LR+, LR–, and AUC with 95% CIs. $P<.05$ was considered statistically significant. All analyses were conducted using SPSS Statistics version 23.0 (IBM, Corp., Armonk, NY) and Stata version 14.0 (Stata Corp., College Station, TX).

Results

Figure 1 shows the flowchart of the study population. During the study period, 2,827 individuals aged 70 and older visited the LUMC, Alrijne Hospital, and HMC Bronovo EDs, of whom 2,554 were eligible, of whom 83.4% (2,130/2,554) were included. For the present study, all individuals discharged from the ED to home were used (51.3%, 1,093/2,130).

The baseline characteristics of the study population are shown in Table 1. Median age was 79 (IQR 74–84), 486 participants (44.5%) were male, and 438 (40.1%) arrived by ambulance. Median Katz ADL score was 0 (IQR 0–1), and 189 participants (18.2%) had cognitive impairment.

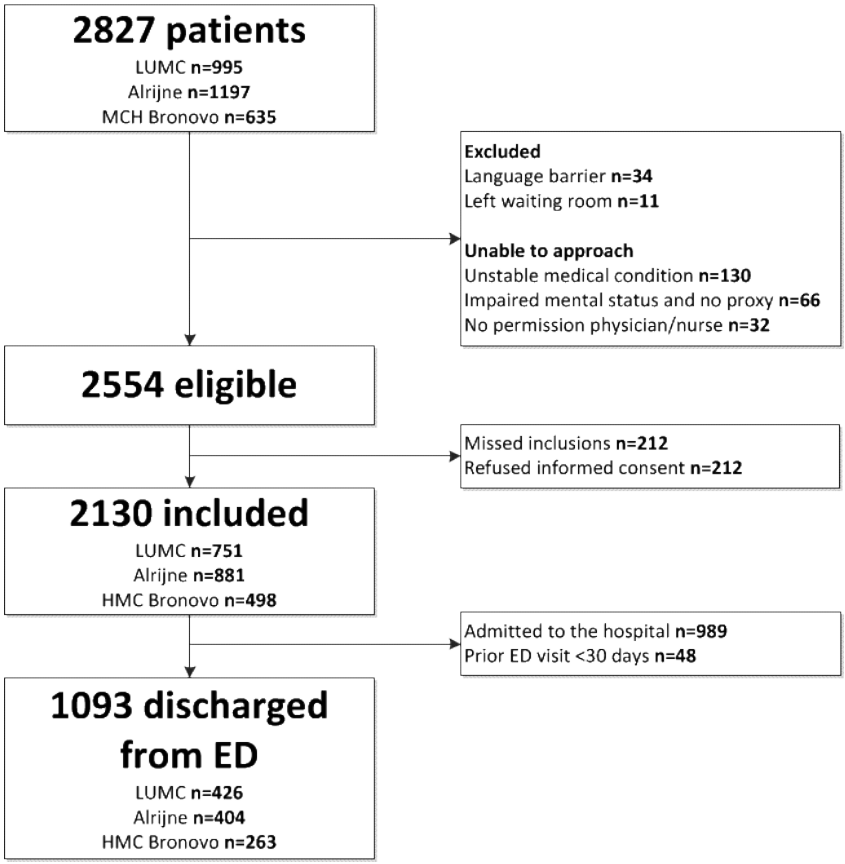

Figure 1: Flowchart of study population

Predictors of an ED revisit

One hundred fourteen of 1,093 participants (10.4%) had an ED revisit within 30 days of follow-up (Supplementary Figure S1): 33 from the LUMC, 48 from Alrijne Hospital, and 33 from HMC Bronovo. Baseline characteristics of those with and without an ED revisit can be found in Supplementary Table S2. In the multivariable analysis, an increase in age of 1 year above the age of 70 (hazard ratio (HR)=0.96, 95% CI=0.92–0.99), male sex (HR=1.61, 95% CI=1.05–2.45), polypharmacy (HR=2.06, 95% CI=1.34–3.16), and cognitive impairment (HR=1.71, 95% CI=1.02–2.88) were associated with a 30-day revisit (Table 2). A prediction model was developed that included age, sex, polypharmacy, and impaired cognition (Supplementary Table S3). Accuracy of the prediction model was poor, with an AUC of (95% CI=0.60–0.70), the Hosmer and Lemeshow goodness of fit p-value was .86, and the formula to compute the individual risk of a revisit can be found in Supplementary Table S3 footnotes. The positive predictive value was 0.18 (95% CI=0.13–0.24) for the 20% of patients at highest risk (Supplementary Table S4).

Table 1: Baseline characteristics of older patients discharged from the emergency department

	N = 1093
Demographics	
Age (years), median (IQR)	79 (74-84)
Male, n(%)	486 (44.5%)
Living arrangement, n (%)	
Independent with others	603 (55.2%)
Independent alone	396 (36.2%)
Residential care/nursing home	94 (8.6%)
High education, n (%)	250 (23.0%)
Severity of medical condition	
Arrival by ambulance, n(%)	438 (40.1%)
Triage urgency, n(%)	
> 1 hour (green)	434 (39.7%)
< 1 hour (yellow)	543 (49.7%)
< 10 min (orange)	116 (10.6%)
Chief complaint, n(%)	
Minor trauma	463 (42.4%)
Chest pain	206 (18.8%)
Malaise	141 (12.9%)
Abdominal pain	75 (6.9%)
Dyspnea	68 (6.2%)
Loss of consciousness	52 (4.8%)
Other	88 (8.1%)
Fall prior to ED visit, n (%)	352 (32.2%)
Geriatric parameters	
Polypharmacy, n(%)	560 (51.2%)
Use of walking device, n(%)	406 (37.3%)
Katz ADL score, median (IQR)	0 (0-1)
Cognitive impairment, n(%)^	189 (18.2%)

N = Number, IQR = inter quartile range, ED = Emergency Department, ADL = Activities of Daily Living
Missing data: 7 level of education, 5 use of walking device, 13 Katz ADL score, 53 cognition scores
^Cognitive impairment includes patients with a 6 CIT score ≥11 and includes 69 patients with a history of dementia

Table 2: Univariate and multivariate association of baseline characteristics with 30-day ED revisits in older patients discharged from the emergency department

	Multivariable HR (95%CI)*	**p-value**
Demographics		
Age (years), median (IQR)	0.96 (0.92-0.99)	0.014
Male, n(%)	1.61 (1.05-2.45)	0.027
Living arrangement, n (%)		
Independent with others	Ref	
Independent alone	1.07 (0.69-1.66)	0.749
Residential care/nursing home	0.78 (0.32-1.91)	0.590
High education, n (%)	0.99 (0.61-1.61)	0.975
Hospital, n (%)		
LUMC	Ref	
Alrijne	1.85 (1.12-3.04)	0.016
HMC Bronovo	2.23 (1.32-3.79)	0.003
Severity of medical condition		
Arrival by ambulance	1.12 (0.72-1.74)	0.609
Triage urgency		
> 1 hour (green)	Ref	
< 1 hour (yellow)	1.05 (0.69-1.62)	0.811
< 10 min (orange)	1.04 (0.47-2.33)	0.918
Chief complaint		
Minor trauma	Ref	
Chest pain	0.60 (0.29-1.22)	0.195
Malaise	0.68 (0.32-1.43)	0.310
Abdominal pain	1.16 (0.53-2.58)	0.708
Dyspnea	1.47 (0.72-3.02)	0.291
Loss of consciousness	0.55 (0.16-1.88)	0.344
Other	1.25 (0.62-2.51)	0.532
Fall prior to ED visit	0.80 (0.45-1.41)	0.438
Geriatric parameters		
Polypharmacy	2.06 (1.34-3.16)	0.001
Use of walking device	1.32 (0.81-2.15)	0.265
Katz ADL score	0.96 (0.80-1.17)	0.709
Cognitive impairment	1.71 (1.02-2.88)	0.044

HR = Hazard ratio, 95%CI = 95% confidence interval, ED = emergency department, ADL = activities of daily living, ref = reference

*All listed variables were included. The multivariable analysis includes 1020 patients with 107 events.

The association between ED revisits and outcomes

Fifty-one of 1,093 participants (4.7%) were lost to follow-up for data on physical functioning, but we verified that they were alive from municipal records. These individuals were considered not to be experiencing functional decline. After allocation based on 1:1 propensity score matching, baseline covariates were similar in the matched cohort (Supplementary Table S5). Patients with an early ED revisit were at higher risk of experiencing the composite outcome (OR=1.99, 95% CI=1.06–3.71, p=.03) (Table 3). The predictive performance of an early ED revisit had a sensitivity of 0.63 (95% CI=0.49–0.75), specificity of 0.54 (95% CI=0.46–0.62), LR+ of 1.37 (95% CI=1.05–1.78), LR– of 0.69 (95% CI=0.49–0.98), and AUC of 0.59 (95% CI=0.50–0.67).

Table 3: The association of an early revisit with functional decline or mortality in the propensity score matched cohort

	Total cohort N = 1093	**Early revisits N = 108**	**Matched controls N = 108**
Composite outcome (%)	253 (23.1%)	35 (32.4%)	21 (19.4%)
OR (95%CI)		1.99 (1.06-3.71)	

OR = odds ratio; 95%CI = 95% confidence interval
A propensity score was not computed for 6 patients (5.3%) with an early ED revisit and 54 patients (5.5%) with no early ED revisit because of missing predictors. The missing predictors are described in legend of *table 1*.

Discussion

Approximately 10% of older adults return to the ED within 30 days after being discharged. Older age, male sex, polypharmacy, and impaired cognition were identified as independent determinants of an ED revisit, but the prediction model is not efficient for clinical use. More importantly, participants with an early ED revisit were twice as likely to experience functional decline or mortality.

It is difficult for ED providers to determine who will return to the ED. To prevent a revisit, identification of older adults who are at high risk enables individualization of discharge planning with concurrent use of the community care network. Prospective studies on 30-day ED revisits in older adults discharged from the ED have been performed.[85, 90, 100, 101] All studies consistently had poor predictive accuracy for ED revisits, with different predictors in the final prediction models. A study in Australia showed that chronic obstructive pulmonary disease, cognitive impairment, a previous ED visit, and a low triage score were independent determinants of a revisit,[85] whereas a study in Canada identified history of diabetes, a recent ED visit, hospital admission during the previous 6 months, depression, and lack of support as independent determinants of a revisit.[90] A risk nomogram on 30-day ED revisits was developed in Austria and included prior number

of attendances, age, male sex, polypharmacy, cognition score, and malignancy,[83] but discrimination of the model was poor, and the ED revisit rate of 28% was much high than in our study.[102] We were unable to replicate all findings of previous studies, possibly because we have looked to other predictors than the other studies. Only age, male sex, polypharmacy, and cognitive impairment were associated with risk of an ED revisit. In line with previous studies, we were unable to develop an accurate prediction model on 30-day ED revisits. Other factors that may influence the risk of returning to the ED may explained the differences between studies in predictors associated with an ED revisit and the difficulties of developing a useful prediction model. Four common themes are described that are related to a revisit; the individual, severity of illness, organization of the healthcare system, and the role of the clinician.[103] Within these themes, some factors take too much time or are difficult to measure in the ED, such as individuals› personal experiences and feelings after being discharged, such as fear, uncertainty, and lack of trust in the system.[104] Other factors can potentially positively be influenced, like aftercare provided by the general practitioner and various home health and paramedicine initiatives. To prevent a revisit, randomized controlled trials and cohort studies were conducted,[83, 84, 87, 105, 106] but the number of ED revisits could not be reduced.[107] It may be that the intervention is inefficient, but it is more likely that the intervention is ineffective in those in whom it is used. The mixed results and heterogeneous causes of revisits complicate identification of older adlts who will return to the ED and challenges whether ED-based interventions can efficiently and reliably reduce ED revisits.

To our knowledge, the association between early ED revisits and subsequent adverse health outcomes has never been evaluated, despite the fact that it is used as a quality-of-care indicator.[103] After balancing the population on the propensity of experiencing an early ED revisit, participants with an early ED revisit were twice as likely to experience functional decline or mortality 3 months after the index visit. Although the difference was not significant, participants with an ED revisit used more medications and were more likely to have cognitive impairment than matched controls. It may be that adverse drug events resulted in ED revisits or cognitive impairment played a role in the reason to return to the ED or it may be proxies for illness severity. Alternatively, participants with an ED revisit were comparable in most characteristics with those without an ED revisit. It could be that an early ED revisit incorporates multiple determinants that are difficult to quantify but are associated with adverse health outcomes, such as the ability to rely on a caregiver, adherence to discharge instructions, and duration of recovery. Early ED revisits are considered to be a negative outcome for older adults, although an ED revisit could be part of the treatment plan (e.g., return to ED when symptoms worsen while being treated with oral antibiotics). Although any visit to the hospital is a burden, other outcomes such as functional dependency are at least as important for older adults. Based on our results,

an early ED revisit can be considered as a predictor of functional decline or mortality, even in older adults that are not presenting in extremis.

Clinical implications

Improving quality of care for older adults is among the highest priorities of multiple stakeholders (American Geriatrics Society, Hartford Foundation Society for Academic Emergency Medicine, National Institute on Aging, New Frontiers). It is challenging to assess vulnerability in the ED setting[108] and to discern which individuals would benefit most from specific geriatric follow-up interventions.[109] A decision rule can assist clinicians in making that decision,[73] but existing and widely used screening instruments lack the ability to accurately distinguish who is at the highest and lowest risk.[4] Results of the present study call into question whether it is efficient to screen for individuals at high risk of an early ED revisit routinely. Instead of using early revisit as a negative or nonreimbursable outcome, it could be used as an independent predictor of adverse outcomes in new or existing prediction models. In the clinic, an early revisit should trigger a thorough assessment of other geriatric determinants of adverse outcomes, with the aim of positively influencing outcomes.[110]

Our study has several limitations. First, we were not able to investigate all possible predictors associated with ED revisits or adverse outcomes, such as medical history, presence of caregivers, and visits to a general practitioner shortly after discharge. The study was designed to reflect daily practice as much as possible, and therefore the aim was to include a representative percentage of patients rather than a subset of the ED population who were healthy enough to complete all questionnaires. As a consequence, the number of potential confounding variables assessed was reduced. Second, it is possible that we missed ED revisits to other hospitals, although we proactively asked participants over the telephone and cross-referenced participants. Third, selection bias could have occurred, because 4.7% of participants were lost to follow-up. We verified in municipal records that these individuals were alive. Instead of excluding them from further analysis, we considered them as not having functional decline, which could have resulted in an underestimation of the outcome. In a sensitivity, results were comparable when these participants were excluded or assumed the have functionally declined. Fourth, the power to develop the prediction model for ED revisits was not sufficient. Results were similar when repeating the analysis with fewer candidate predictors by excluding chief complaint and living arrangement. A strength of the present study was that an unselected representative group of older adults (83% of all eligible individuals) attending the ED was included and that the follow-up rate was high (95.3%). Second, because of the propensity score matching analysis, participants and matched controls had a comparable chance for an early ED revisit.

In conclusion, older age, male sex, polypharmacy, and cognitive impairment were independently associated with a 30-day ED revisit. No useful clinical prediction model could be developed. However, an early ED revisit is a strong new predictor of adverse outcomes in older adults.

References

1. Arendts, G., et al., *Derivation of a nomogram to estimate probability of revisit in at-risk older adults discharged from the emergency department.* Intern Emerg Med, 2013. 8 (3): p. 249-54.
2. Caplan, G.A., et al., *A randomized, controlled trial of comprehensive geriatric assessment and multidisciplinary intervention after discharge of elderly from the emergency department--the DEED II study.* J Am Geriatr Soc, 2004. 52 (9): p. 1417-23.
3. Lowthian, J., et al., *Unplanned early return to the emergency department by older patients: the Safe Elderly Emergency Department Discharge (SEED) project.* Age Ageing, 2016.
4. McCusker, J., et al., *Hospital characteristics and emergency department care of older patients are associated with return visits.* Acad Emerg Med, 2007. 14 (5): p. 426-33.
5. Mion, L.C., et al., *Case finding and referral model for emergency department elders: a randomized clinical trial.* Ann Emerg Med, 2003. 41 (1): p. 57-68.
6. Costa, A.P., et al., *Geriatric syndromes predict postdischarge outcomes among older emergency department patients: findings from the interRAI Multinational Emergency Department Study.* Acad Emerg Med, 2014. 21 (4): p. 422-33.
7. Deschodt, M., et al., *Characteristics of older adults admitted to the emergency department (ED) and their risk factors for ED readmission based on comprehensive geriatric assessment: a prospective cohort study.* BMC Geriatr, 2015. (15): p. 54.
8. McCusker, J., et al., *Return to the emergency department among elders: patterns and predictors.* Acad.Emerg.Med., 2000. 7 (3): p. 249-259.
9. Banerjee, J., S. Conroy, and M.W. Cooke, *Quality care for older people with urgent and emergency care needs in UK emergency departments.* Emerg Med J, 2013. 30 (9): p. 699-700.
10. Carpenter, C.R., et al., *Optimal older adult emergency care: introducing multidisciplinary geriatric emergency department guidelines from the American College of Emergency Physicians, American Geriatrics Society, Emergency Nurses Association, and Society for Academic Emergency Medicine.* J Am Geriatr Soc, 2014. 62 (7): p. 1360-3.
11. Aminzadeh, F. and W.B. Dalziel, *Older adults in the emergency department: a systematic review of patterns of use, adverse outcomes, and effectiveness of interventions.* Ann.Emerg.Med., 2002. 39 (3): p. 238-247.
12. de Gelder, J., et al., *Predicting adverse health outcomes in older emergency department patients: the APOP study.* Neth J Med, 2016. 74 (8): p. 342-352.
13. Mackway-Jones, K., *Manchester Triage Group.* Emergency Triage, 1997.
14. von Elm, E., et al., *The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies.* Ann Intern Med, 2007. 147 (8): p. 573-7.
15. Moons, K.G., et al., *New Guideline for the Reporting of Studies Developing, Validating, or Updating a Multivariable Clinical Prediction Model: The TRIPOD Statement.* Adv Anat Pathol, 2015. 22 (5): p. 303-5.

16. Landrum, L. and S. Weinrich, *Readmission data for outcomes measurement: identifying and strengthening the empirical base.* Qual Manag Health Care, 2006. 15 (2): p. 83-95.
17. McHugh, M., et al., *More patients are triaged using the Emergency Severity Index than any other triage acuity system in the United States.* Acad Emerg Med, 2012. 19 (1): p. 106-9.
18. Katz, S., et al., *Studies of Illness in the Aged. The Index of Adl: A Standardized Measure of Biological and Psychosocial Function.* JAMA, 1963. (185): p. 914-9.
19. Katzman R, B.T., Fuld P, Peck A, Schechter R, Schimmel H., *Validation of a short Orientation-Memory-Concentration Test of cognitive impairment.* Am J Psychiatry., 1983. Jun (140 (6)): p. 734-9.
20. Tuijl, J.P., et al., *Screening for cognitive impairment in older general hospital patients: comparison of the Six-Item Cognitive Impairment Test with the Mini-Mental State Examination.* Int J Geriatr Psychiatry, 2012. 27 (7): p. 755-62.
21. O'Sullivan, D., et al., *Validation of the 6-Item Cognitive Impairment Test and the 4AT test for combined delirium and dementia screening in older Emergency Department attendees.* Age Ageing, 2018. 47 (1): p. 61-68.
22. Lucke, J.A., et al., *Impaired cognition is associated with adverse outcome in older patients in the Emergency Department; the Acutely Presenting Older Patients (APOP) study.* Age Ageing, 2018. 47 (5): p. 679-684.
23. Steyerberg, E.W., *Clinical prediction models: a practical approach to development, validation, and updating.*, ed. M.G.K.K.J.S.A.T. W.Wong. 2009, New York: Springer.
24. Hosmer DW , L.S., *Applied logistic regression*. 2nd ed. 2000, New York: Wiley.
25. Senn, S., E. Graf, and A. Caputo, *Stratification for the propensity score compared with linear regression techniques to assess the effect of treatment or exposure.* Stat Med, 2007. 26 (30): p. 5529-44.
26. Hastings, S.N., et al., *Frailty predicts some but not all adverse outcomes in older adults discharged from the emergency department.* J Am Geriatr Soc, 2008. 56 (9): p. 1651-7.
27. LaMantia, M.A., et al., *Predicting hospital admission and returns to the emergency department for elderly patients.* Acad Emerg Med, 2010. 17 (3): p. 252-9.
28. Arendts, G., et al., *Use of a risk nomogram to predict emergency department reattendance in older people after discharge: a validation study.* Intern Emerg Med, 2015. 10 (4): p. 481-7.
29. Trivedy, C.R. and M.W. Cooke, *Unscheduled return visits (URV) in adults to the emergency department (ED): a rapid evidence assessment policy review.* Emerg Med J, 2015. 32 (4): p. 324-9.
30. Rising, K.L., et al., *Return visits to the emergency department: the patient perspective.* Ann Emerg Med, 2015. 65 (4): p. 377-386 e3.
31. Guttman, A., et al., *An emergency department-based nurse discharge coordinator for elder patients: does it make a difference?* Acad Emerg Med, 2004. 11 (12): p. 1318-27.
32. Biese, K.J., et al., *Telephone Follow-Up for Older Adults Discharged to Home from the Emergency Department: A Pragmatic Randomized Controlled Trial.* J Am Geriatr Soc, 2018. 66 (3): p. 452-458.
33. Lowthian, J.A., et al., *Discharging older patients from the emergency department effectively: a systematic review and meta-analysis.* Age Ageing, 2015. 44 (5): p. 761-70.

34. Carpenter, C.R. and M. Emond, *Pragmatic barriers to assessing post-emergency department vulnerability for poor outcomes in an ageing society.* Neth J Med, 2016. 74 (8): p. 327-329.
35. Hwang, U. and C. Carpenter, *Assessing geriatric vulnerability for post emergency department adverse outcomes: challenges abound while progress is slow.* Emerg Med J, 2016. 33 (1): p. 2-3.
36. Stiell, I.G. and G.A. Wells, *Methodologic standards for the development of clinical decision rules in emergency medicine.* Ann Emerg Med, 1999. 33 (4): p. 437-47.
37. Carpenter, C.R., et al., *Risk factors and screening instruments to predict adverse outcomes for undifferentiated older emergency department patients: a systematic review and meta-analysis.* Acad Emerg Med, 2015. 22 (1): p. 1-21.
38. Conroy, S.P., et al., *A controlled evaluation of comprehensive geriatric assessment in the emergency department: the 'Emergency Frailty Unit'.* Age Ageing, 2013.

CHAPTER 7

General discussion

General discussion

In the present thesis a screening instrument to identify older patients at risk for adverse outcomes was developed, validated, refined and cross-validated: the APOP screener. The results of these give rise to initiate implementation studies of the APOP screener for prediction of adverse health outcomes in older patients visiting the ED.

Clinical applicability

In the hospital, the ED is a central access point for acute care and the care provided 'set the stage' for subsequent care. Physicians regularly have to make diagnostic, therapeutic and management decisions based on limited available information.[1] The gold standard to assess older patients' needs across multiple domains is a comprehensive geriatric assessment(CGA).[2] Although a CGA can improve outcomes,[3, 4] it takes too much time to conduct systematically in all patients in an environment where time is scarce and where the condition of the patient is unstable. The present results show that the newly developed APOP screener could be used in the ED as a first step to identify patients at highest risk of adverse outcomes. The next step is to initiate interventions for the high risk patients. Some caution is warranted for the final implementation, because it is known that of all developed prognostic screeners only few will finally be used in practice.[5] To increase the chance of successful implementation, multiple professionals were involved in the development, validation and refinement phases. Currently we are conducting an implementation study to test experiences of the APOP screener in clinical practice and solve a wide range of barriers.[6] Going through all these stages could finally result into a widely accepted screener.

Practice based evidence

Practice based evidence, a procedure for gathering good-quality data from routine practice,[7] is particularly suitable for the heterogeneous older patients populating presenting on the ED.[8] In our study design, available data from the clinic and relevant patient characteristics were used for statistical analysis.[7] An important feature of practice based evidence is a representative study population. The use of minimal inclusion criteria can ensure high generalizability and high external validity. In the APOP study all older ED patients, with exception of critically ill patients, were eligible. Patients were included in four different hospitals, both on daytime hours and in the evening and both on week and weekend days. A total of 3.544 individuals visited the ED during our study period of which 2.629 (74.2%) were finally included. As a comparison, during enrolment period of the development study of the ISAR screening instrument a total 4.157 individuals visited the ED (daytime hours and week days), of which 1.885 (45%) were included.[9] For the TRST screening instrument a total of 7.017 individuals visited the ED during the study period of which 650 were included (9.2%).[10] The representative study population

and practice based data enabled to develop a screener which is suitable to use for older patients visiting the ED.

Presentation formats

Since the ED is a multidisciplinary setting, it is of importance that the screener will be accepted by the various disciplines involved in the treatment of older patients. Choices, which have to be made during the stages towards implementation, can affect the success. One important choice is the format of presentation, in which two formats can be distinguished. The first way of presenting is prediction models, such as regression formula's, nomograms and prognosis programs. The second is decision rules, such as score chart rules and regression trees.[11] In a prediction model, characteristics are combined to provide an absolute risk of experiencing the outcome and in a decision rule a subjective chosen cut-off point will generate advice. Existing implemented screening instruments, such as the ISAR and TRST, are used as bedside score chart rules, a format of a decision rule. They are easy to understand and the cut-off point will determine whether the patients is at 'high-risk'. However, the chosen cut-off is subjective and could influence the success of follow-up interventions, as described in chapter 4. The APOP screener is both a prediction model and a decision rule (http://screener.apop.eu, figure 1). The regression formula automatically calculates the absolute risk of experiencing the outcome by using the application and a subjective cut-off is chosen in which patients are at 'high risk' for functional decline or mortality. By presenting absolute risks, different cut-offs can be used in different settings. We decided to set the cut-off so that on average 1 out of 5 patients will be considered at high risk. This way relatively a low percentage of patients will incorrectly considered as high risk and there is sufficient capacity to conduct follow-up actions or interventions in all high risk patients.

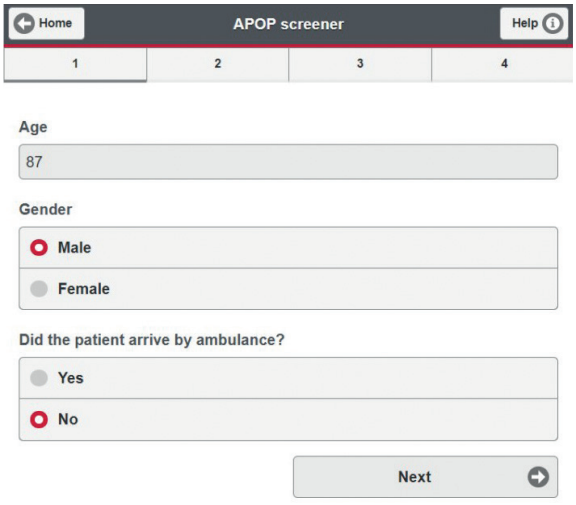

Figure 1: Screenshot of the APOP screener

Selection of predictors

Development of a prediction model includes the selection of predictors.[12] In the present thesis different techniques were used to select the final predictors of a prediction model. In chapter 2 restricted cubic splines techniques were applied for optimal use of continues variables. In everyday clinical practice we consider continuous laboratory results or vital measurements normal or abnormal based on a predefined cut-off. As an example, systolic blood pressures of 83, 162 and 190 millimetre of mercury (mmHG) are all considered as abnormal. It is likely that data will show a strong association with adverse outcomes in patients with a blood pressure of 83 mmHg, no association with a blood pressure of 162 mmHg and a weak association with a blood pressure of 190 mmHg. Restricted cubic splines can model the non-linear associations in the development phase of a prediction model. Usable categories can subsequently be determined in a later phase. In chapter 3 the prediction model was derived with binary logistic regression via backward elimination. As a result, the number of candidate predictors was limited, resulting in the prediction model. The above mentioned techniques will result in the best possible statistical model, but do not take other factors of success into account, such as the effort or time needed to ask questions. From a clinical point of view, the optimal prediction model is a balance between statistical performance and optimal use in clinic. In chapter 5 another strategy was therefore used to redevelop the APOP screener. Five major criteria were formulated for selection of predictors first. Subsequently, in a multidisciplinary meeting consensus was reached for the final screener. The model was robust by validation via the leave-one-hospital out procedure, but external validation in new datasets is still needed to test external validity. In summary, many statistical techniques are available to select the final predictors, but we prefer to select the predictors based on predefined criteria to increase the chance of successful implementation.

Accuracy and predictive values

New prediction models for older patients visiting the ED are popping up frequently in journals, and there will always be discussion which model is the best to use in a specific setting. In addition to representativeness and the format of presentation, also measures of accuracy and performance can be weighed. The sensitivity and specificity are commonly used for interpretation and can be used to help ruling in or ruling out patients who will experience the adverse outcome, such as functional decline or mortality.[13] The area under the receiver operating characteristic curve (AUC) is a summary measure of correct classification of patients who do and don't experience the adverse outcome.[14] The likelihood ratio (LR) of a test is seen as one of the best measures of accuracy and quantifies the change in probability of the adverse outcome based on the screening result.[13, 15] In most settings, screening results are set to rule out patients who will not experience the adverse outcome, which implies a high sensitivity.[16] As a drawback relatively many patients are considered falsely at high risk (low specificity). I believe that the APOP screener

should be 'clinical sensible' rather than having a high sensitivity. The consequence of missing older patients who will experience the adverse outcome should outweigh the benefits of selection of a smaller group of patients really at higher risk. The APOP screener cut-off was therefore set to ensure a high specificity, which resulted in a higher positive predictive value (PPV). This way intensive and costly resources are targeted in patients with a higher chance to benefit from. The harm to low-risk patients will be minimal by providing usual care. Taken all together, by using the APOP screener the first step has been taken to identify older emergency patients at highest risk for adverse outcomes.

Future perspectives

The present thesis describes a variety of choices which have been made and steps which have been taken to develop and validate the APOP screener. It is only the beginning. The impact of implementation of the APOP screener is unknown and important to evaluate.[17] Is it possible to integrate screening at the ED in existing care? Does it direct physicians' decisions? Currently we are conducting an implementation study in the LUMC to test feasibility and impact of the APOP screener in daily practice. By using the RE-AIM framework we expect to improve the sustainable adoption and implementation of the screener.[18]

Not only the APOP screener, but also the in-hospital and out-of-hospital follow-up interventions in high-risk patients need to implemented. Future research should focus on the interventions. Which interventions decrease functional outcomes? What is the patient perspective of the APOP screening program? How to cooperate with other health care professionals in the acute setting, such as the general practitioners?

Technology will continue to develop. In theory every observation, measurement or test adds information and could influence prediction. It is fascinating to imagine what we can do by using big data in prognostication and how care will then be organized, including the role of the human beings in white coats.

Clinical implications

The APOP screener is now available, which will lead to a number of discussions. First, is the APOP screener the best screener to identify older patients at risk for adverse outcomes? In the Dutch ED's it possibly is. In ED's of other countries generalizability will be affected by case-mix differences, what can be tested in validation studies and could be solved by updating the model. One universal perfect screener will simply not exist. Second, what to do with the patients considered at high risk? Depending on the possibilities within the setting and the severity of the acute illness, small to intensive interventions could be executed. Depending on the findings a subsequent geriatric assessment can be performed, for example during hospital admission by a geriatric consultation team,

at the outpatient clinic or via the general practitioner. The multidimensional individual approach should ensure that we do not lag behind incidents and improve outcomes on longer term. Third, what will the screener adds comparing to current care? Older patients are treated by multiple health care professionals and it is difficult to systematically take their vulnerability into account. A 'clinical sensible' screener could assist the professional and would be a way to increase awareness in general. Moreover, it could contribute to an unambiguous way of communication. Fourth, who is responsible for the older patients in the ED? We all are. Implementation of the APOP screener should go hand in hand with training sessions for all the personnel on the ED. The EDs should adapt the organisation of care to older patients within their own context. The patient will still be treated by the assigned specialism, but we should all be aware of the possible adverse outcomes on longer term for some patients, independent of the acute illness at that moment.

References

1. Croskerry, P. and G. Norman, *Overconfidence in clinical decision making.* Am J Med, 2008. 121 (5 Suppl): p. S24-9.
2. Rubenstein, L.Z., et al., *Impacts of geriatric evaluation and management programs on defined outcomes: overview of the evidence.* J Am Geriatr Soc, 1991. 39 (9 Pt 2): p. 8S-16S; discussion 17S-18S.
3. Conroy, S.P., et al., *A controlled evaluation of comprehensive geriatric assessment in the emergency department: the 'Emergency Frailty Unit'.* Age Ageing, 2013.
4. Ellis, G., T. Marshall, and C. Ritchie, *Comprehensive geriatric assessment in the emergency department.* Clin Interv Aging, 2014. (9): p. 2033-43.
5. Moons, K.G., et al., *Prognosis and prognostic research: application and impact of prognostic models in clinical practice.* BMJ, 2009. (338): p. b606.
6. Peters, D.H., et al., *Implementation research: what it is and how to do it.* BMJ, 2013. 347: p. f6753.
7. Margison, F.R., et al., *Measurement and psychotherapy. Evidence-based practice and practice-based evidence.* Br J Psychiatry, 2000. (177): p. 123-30.
8. Aminzadeh, F. and W.B. Dalziel, *Older adults in the emergency department: a systematic review of patterns of use, adverse outcomes, and effectiveness of interventions.* Ann.Emerg.Med., 2002. 39 (3): p. 238-247.
9. McCusker, J., et al., *Detection of older people at increased risk of adverse health outcomes after an emergency visit: the ISAR screening tool.* J.Am.Geriatr.Soc., 1999. 47 (10): p. 1229-1237.
10. Meldon, S.W., et al., *A brief risk-stratification tool to predict repeat emergency department visits and hospitalizations in older patients discharged from the emergency department.* Acad Emerg Med, 2003. 10 (3): p. 224-32.
11. Steyerberg, E.W., *Clinical prediction models: a practical approach to development, validation, and updating.*, ed. M.G.K.K.J.S.A.T. W.Wong. 2009, New York: Springer.
12. Steyerberg, E.W. and Y. Vergouwe, *Towards better clinical prediction models: seven steps for development and an ABCD for validation.* Eur Heart J, 2014. 35 (29): p. 1925-31.
13. Drobatz, K.J., *Measures of accuracy and performance of diagnostic tests.* J Vet Cardiol, 2009. 11 (Suppl 1): p. S33-40.
14. Zweig, M.H. and G. Campbell, *Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine.* Clin Chem, 1993. 39 (4): p. 561-77.
15. McGee, S., *Simplifying likelihood ratios.* J Gen Intern Med, 2002. 17 (8): p. 646-9.
16. Stiell, I.G. and G.A. Wells, *Methodologic standards for the development of clinical decision rules in emergency medicine.* Ann Emerg Med, 1999. 33 (4): p. 437-47.
17. Toll, D.B., et al., *Validation, updating and impact of clinical prediction rules: a review.* J Clin Epidemiol, 2008. 61 (11): p. 1085-94.
18. Glasgow, R.E., T.M. Vogt, and S.M. Boles, *Evaluating the public health impact of health promotion interventions: the RE-AIM framework.* Am J Public Health, 1999. 89 (9): p. 1322-7.

CHAPTER 8

English summary
Nederlandse samenvatting
List of co-authors
List of publications
Curriculum vitae
Dankwoord

English summary

Introduction

The rate of adverse health outcomes 90 days after an emergency department (ED) visit is high for older patients. A comprehensive geriatric assessment (CGA) can effectively identify patients at highest risk and improve outcomes. Unfortunately, the fast-paced environment and the medical condition of the acutely ill older patient make it virtually impossible to perform a CGA in all patients in the ED. A two-step approach could be useful in order to improve adverse health outcomes. Older patients at highest risk need to be identified first, subsequently targeted interventions can be conducted. The aim of the present thesis was to develop and validate a screening instrument to predict adverse health outcomes in older patients visiting the ED.

Summary of the chapters

Chapter 2 describes the characteristics and outcomes of acutely admitted older patients. It shows that approximately 1 out of 5 older patients deceased 90 days after acute hospital admission. A prediction model containing information about illness severity and geriatric parameters was developed to predict 90-day mortality. Using this model, 53% of the patients in the highest risk decile were deceased within 90 days. In **chapter 3** a new prediction model for predicting 90-day functional decline or mortality was developed and validated for older patients visiting the ED: the APOP (Acutely Presenting Older Patient) screener. The prediction model included early available predictors, to ensure measures can be conducted in an early phase of presentation. We showed that the existing Identification of Seniors At Risk (ISAR) screening tool is more useful to 'rule out' patients at high risk and that the APOP screener was more useful to identify those at highest risk. **Chapter 4** aims to optimize the predictive value of the Identification of Seniors at Risk – Hospitalized Patients (ISAR-HP) by using different cut-off points in acutely hospitalized older patients. The ISAR-HP classifies a high percentage of patients incorrectly as high risk and performance improved marginally by raising the cut-off.

Many developed screening instruments to predict adverse health outcomes in older patients visiting the ED are frequently rejected in clinical practice. In **chapter 5** the APOP screener was refined according to international methodological standards in order to increase the chance for successful implementation. The predictors included in the APOP screener were chosen based on predefined criteria. Performance of the screener was satisfactory, cross-validation between four hospitals was successful and a pilot study among triage nurses showed that it was feasible to complete screening during triage. In **chapter 6** the predictors of ED revisits and the association of an ED revisit with 90-day functional decline or mortality were studied. This chapter shows that a few independent predictors were identified, but no useful clinical prediction model could be developed.

More importantly, an ED revisit was strongly associated with functional decline and mortality and can possible be used as new predictor in future prediction models.

Chapter 7 includes the key findings of this thesis. In the development phase of the APOP screener choices have been made in order to make a useful screener for clinical practice. A representative study population was included by using minimal inclusion criteria and including patients 7 days a week. A wide range of disciplines were involved in the research project. In order to have a balance between accuracy and optimal use in practice the APOP screener was redeveloped based on pre-defined criteria in a multi-disciplinary meeting. The result of the APOP screening shows an absolute risk of the adverse health outcome and classifies a group of patients at high risk. In contrast to comparable existing screening instruments, the cut-off was set to ensure high specificity. This way relatively a low percentage of older patients will be incorrectly classified as high risk and (costly) follow-up interventions will be targeted in older patients with a higher chance to benefit from.

The APOP screener has been developed and recently an implementation study was conducted to test feasibility and impact of the screener. Future studies should focus on the effect of (single) interventions on outcomes, examine the patient perspective on the APOP screening program and investigate how to cooperate with other health care professionals in the acute setting, such as the general practitioners.

Nederlandse samenvatting

Introductie

De snelheid van functionele achteruitgang en overlijden 90 dagen na een spoedeisende hulp (SEH) bezoek is hoog voor oudere patiënten. Met een "comprehensive geriatric assessment" (CGA), een uitgebreide geriatrische beoordeling, kunnen patiënten met een verhoogde kans op achteruitgang geïdentificeerd worden en vervolgens stappen genomen worden om de uitkomsten te verbeteren. Door de tijdsdruk en de medische toestand van de acuut zieke oudere patiënt is het helaas onmogelijk om een CGA uit te voeren bij alle oudere patiënten op de SEH. In deze setting zou een stapsgewijze aanpak een passende oplossing kunnen zijn om gezondheidsuitkomsten te verbeteren. In de eerste stap zal met behulp van een korte screening een voorselectie gemaakt moeten worden van oudere patienten met het hoogste risico op functionele achteruitgang. In de tweede stap kunnen voor deze hoog-risico patiënten maatregelen genomen worden om de uitkomsten te verbeteren. Het doel van dit proefschrift was om een screeningsinstrument te ontwikkelen en te valideren om functionele achteruitgang te voorspellen bij ouderen patienten die de SEH bezoeken.

Samenvatting van de hoofdstukken

Hoofdstuk 2 beschrijft de kenmerken en uitkomsten van acuut opgenomen oudere patiënten in het ziekenhuis. De resultaten tonen aan dat ongeveer 1 op de 5 oudere patiënten 90 dagen na acute ziekenhuisopname overlijdt. Een screeningsinstrument met informatie over de ernst van de ziekte en geriatrische kenmerken werd ontwikkeld om de kans op sterfte binnen 90 dagen te voorspellen. Met dit model kon een groep patiënten worden geïdentificeerd met een verhoogd risico op overlijden, waarvan uiteindelijk 53% binnen 90 dagen daadwerkelijk was overleden. In **hoofdstuk 3** werd een nieuw screeningsinstrument voor het voorspellen van 90 dagen functionele achteruitgang of sterfte ontwikkeld en gevalideerd voor oudere patiënten die de SEH bezoeken: de APOP (Acuut Presenterende Oudere Patiënt) screener. De screener bestaat uit vragen die na binnenkomst afgenomen kunnen worden, om te zorgen dat bij hoog-risico patiënten direct maatregelen genomen kunnen worden. In dit hoofdstuk wordt tevens aangetoond dat de al bestaande "identification of seniors at risk" (ISAR) screener met name bruikbaar is om uit te sluiten dat patiënten functioneel achteruitgaan en dat de APOP screener juist geschikter is om de patiënten met het hoogste risico op achteruitgang te identificeren. **Hoofdstuk 4** is gericht op het optimaliseren van de voorspellende waarde van de "Identification of Seniors at Risk – Hospitalized Patients" (ISAR-HP) screener. De ISAR-HP wijst een groot percentage van de patiënten ten onrechte als hoog-risico aan en de voorspellingen van de ISAR-HP verbeterden marginaal door het afkappunt te veranderen.

Veel van de ontwikkelde screeningsinstrumenten die achteruitgang bij oudere patiënten op de SEH voorspellen, worden in de praktijk om verschillende redenen uiteindelijk niet gebruikt. In **hoofdstuk 5** werd de APOP screener aangepast volgens internationale methodologische standaarden om de kans op een succesvolle implementatie op de werkvloer te vergroten. De definitieve vragen in de APOP screener werden gekozen op basis van vooraf gedefinieerde criteria. De prestaties van de screener waren bevredigend, de kruisvalidatie in vier ziekenhuizen was succesvol en een pilot-studie onder triage verpleegkundigen toonde aan dat het haalbaar was om de screening tijdens de triage te voltooien. In **hoofdstuk 6** werden voorspellers van een herbezoek aan de SEH en de associatie tussen een herbezoek aan de SEH en 90 dagen functionele achteruitgang bestudeerd. Dit hoofdstuk laat zien dat enkele onafhankelijke voorspellers werden geïdentificeerd, maar er kon geen bruikbaar predictief model worden ontwikkeld. Veel belangrijker was de bevinding dat een herbezoek aan de SEH sterk geassocieerd is met functionele achteruitgang en sterfte. Een herbezoek aan de SEH zou mogelijk gebruikt kunnen worden als nieuwe voorspeller in toekomstige screeningsinstrumenten.

Hoofdstuk 7 bediscussieert de belangrijkste bevindingen van dit proefschrift. Een aantal keuzes zijn in de ontwikkelingsfase van de APOP screener gemaakt om een geschikte screener voor de praktijk te ontwikkelen. In de studie opzet is gekozen voor minimale inclusie criteria, waarbij in alle centra op intensieve wijze zowel overdag als 's avonds en in week- en weekenddagen patiënten geïncludeerd zijn. Deze opzet heeft bijgedragen aan een representatieve patiëntenpopulatie. Om de klinische toepasbaarheid te vergroten zijn in de fases van ontwikkeling verschillende disciplines betrokken bij het proces. Uiteindelijk is in een multidisciplinair overleg aan de hand van vooraf gekozen criteria besloten wat de vragen van de definitieve APOP screener moeten zijn. Hoewel op deze manier ingeleverd wordt op accuraatheid van de screener, wordt de kans op succesvolle implementatie in de praktijk vergroot. Het APOP screeningsresultaat laat zowel het absolute risico op achteruitgang zien, als een waarschuwing dat de patiënt tot de hoog-risico groep behoort. In tegenstelling tot bestaande screeningsinstrumenten is gekozen is voor een afkappunt met hoge specificiteit voor de hoog-risico groep. Op deze manier worden relatief weinig patiënten onterecht als hoog-risico aangewezen en kunnen de (kostbare) vervolg maatregelen doelmatig ingezet worden.

De APOP screener is ontwikkeld en recent is een implementatie studie uitgevoerd om de haalbaarheid en de impact van de screener te testen op de werkvloer. Vervolgstudies zouden zich kunnen focussen op de verschillende soorten interventies bij hoog-risico patiënten, zoals het effect van (poliklinische) geriatrische follow-up, telefonische overdracht aan de huisarts of het streven naar vroege mobilisatie. Welke hebben een positief effect op gezondheidsuitkomsten? Wat vinden de patiënten van het APOP screening programma? Hoe zouden we in de acute setting optimaal samen kunnen werken met andere zorgverleners, zoals de huisartsen?

List of co-authors

Anja M. Booijen, Department of Emergency Medicine, Leiden University Medical Center, Leiden, the Netherlands

Anne J. Fogteloo, Department of Internal Medicine, section on Acute Care, Leiden University Medical Center, Leiden, the Netherlands

Anniek Brink, Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands.

Anton J.M. de Craen, Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands. *Deceased 17-01-2016*

Bas de Groot, Department of Emergency Medicine, Leiden University Medical Center, Leiden, the Netherlands

Christian Heringhaus, Department of Emergency Medicine, Leiden University Medical Center, Leiden, the Netherlands

Edith Haenen, Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands

Ewout W. Steyerberg, Department of Medical Statistics, Leiden University Medical Center, Leiden, The Netherlands

Gerard Jan Blauw, Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands

Henk G. Klop, Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands

Jacinta A. Lucke, Department of Emergency Medicine, Leiden University Medical Center, Leiden, The Netherlands

Jelmer Alsma, Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands.

Kanar Mesri, Department of Internal Medicine, section on Acute Care, Leiden University Medical Center, Leiden, the Netherlands

Laura C. Blomaard, Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands

Noor Heim, Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands

Olaf M. Dekkers, Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands

Roelof A.J. Smit, Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands

Sander Anten, Department of Internal Medicine, section on Acute Care, Alrijne Hospital, Leiderdorp, the Netherlands

Shantaily D. Lourens, Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands

Simon P. Mooijaart, Department of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, the Netherlands and Institute of Evidence-Based Medicine in Old Age, Leiden, the Netherlands

Stephanie C.E. Schuit, Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, the Netherlands.

List of publications

Broekhuizen, K., **J. de Gelder**, C. A. Wijsman, L. W. Wijsman, R. G. Westendorp, E. Verhagen, P. E. Slagboom, A. J. de Craen, W. van Mechelen, D. van Heemst, F. van der Ouderaa and S. P. Mooijaart (2016). "An Internet-Based Physical Activity Intervention to Improve Quality of Life of Inactive Older Adults: A Randomized Controlled Trial." J Med Internet Res 18(4): e74.

de Gelder, J., J. A. Lucke, N. Heim, A. J. de Craen, S. D. Lourens, E. W. Steyerberg, B. de Groot, A. J. Fogteloo, G. J. Blauw and S. P. Mooijaart (2016). "Predicting mortality in acutely hospitalized older patients: a retrospective cohort study." Intern Emerg Med 11(4): 587-594.

de Gelder, J., J. A. Lucke, B. de Groot, A. J. Fogteloo, S. Anten, K. Mesri, E. W. Steyerberg, C. Heringhaus, G. J. Blauw and S. P. Mooijaart (2016). "Predicting adverse health outcomes in older emergency department patients: the APOP study." Neth J Med 74(8): 342-352.

Hofman, S. E., J. A. Lucke, N. Heim, **J. de Gelder**, A. J. Fogteloo, C. Heringhaus, B. de Groot, A. J. de Craen, G. J. Blauw and S. P. Mooijaart (2016). "Prediction of 90-day mortality in older patients after discharge from an emergency department: a retrospective follow-up study." BMC Emerg Med 16(1): 26.

de Gelder, J., E. Haenen, J. A. Lucke, H. G. Klop, L. C. Blomaard, R. A. J. Smit, K. Mesri, B. de Groot, A. J. Fogteloo, S. Anten, G. J. Blauw and S. P. Mooijaart (2017). "Optimising the ISAR-HP to screen efficiently for functional decline in older patients." Neth J Med 75(9): 379-385.

de Gelder, J., J. A. Lucke, L. C. Blomaard, A. M. Booijen, A. J. Fogteloo, S. Anten, E. W. Steyerberg, J. Alsma, S. C. E. Klein Nagelvoort Schuit, A. Brink, B. de Groot, G. J. Blauw and S. P. Mooijaart (2018). "Optimization of the APOP screener to predict functional decline or mortality in older emergency department patients: Cross-validation in four prospective cohorts." Exp Gerontol 110: 253-259.

de Gelder, J., J. A. Lucke, B. de Groot, A. J. Fogteloo, S. Anten, C. Heringhaus, O. M. Dekkers, G. J. Blauw and S. P. Mooijaart (2018). "Predictors and Outcomes of Revisits in Older Adults Discharged from the Emergency Department." J Am Geriatr Soc 66(4): 735-741.

Lucke, J. A., **J. De Gelder**, L. C. Blomaard, A. J. Fogteloo, J. Alsma, S. C. E. Schuit, A. Brink, B. De Groot, G. J. Blauw and S. P. Mooijaart (2018). "CAM-ICU may not be the optimal screening tool for early delirium screening in older emergency department patients: a prospective cohort study." Eur J Emerg Med.

Lucke, J. A., **J. de Gelder**, F. Clarijs, C. Heringhaus, A. J. M. de Craen, A. J. Fogteloo, G. J. Blauw, B. Groot and S. P. Mooijaart (2018). "Early prediction of hospital admission for emergency department patients: a comparison between patients younger or older than 70 years." Emerg Med J 35(1): 18-27.

Lucke, J. A., **J. de Gelder**, C. Heringhaus, R. C. van der Mast, A. J. Fogteloo, S. Anten, G. J. Blauw, B. de Groot and S. P. Mooijaart (2018). "Impaired cognition is associated with adverse outcome in older patients in the Emergency Department; the Acutely Presenting Older Patients (APOP) study." Age Ageing 47(5): 679-684.

Lucke, J. A., R. C. van der Mast, **J. de Gelder**, N. Heim, B. de Groot, S. P. Mooijaart and G. J. Blauw (2018). "The Six-Item Cognitive Impairment Test Is Associated with Adverse Outcomes in Acutely Hospitalized Older Patients: A Prospective Cohort Study." Dement Geriatr Cogn Dis Extra 8(2): 259-267.

Van Deudekom, F. J., **J. De Gelder**, J. A. Lucke, A. Oostendorp-Lange, S. Anten, G. J. Blauw, B. De Groot and S. P. Mooijaart (2018). "Determinants of self-rated health in older adults before and 3 months after an emergency department visit: a prospective study." Eur J Emerg Med.

Lucke, J. A., **J. de Gelder**, L. C. Blomaard, C. Heringhaus, J. Alsma, S. C. E. Klein Nagelvoort Schuit, A. Brink, S. Anten, G. J. Blauw, B. de Groot and S. P. Mooijaart (2019). "Vital signs and impaired cognition in older emergency department patients: The APOP study." PLoS One 14(6):

Curriculum Vitae

Jelle de Gelder werd op 4 januari 1987 geboren te Leiderdorp. Na het behalen van het gymnasiumdiploma aan het Stedelijk Gymnasium Leiden in 2005, begon hij met de studie biomedische wetenschappen aan de Vrije Universiteit Amsterdam. In 2006 is hij gestart met de studie geneeskunde aan de Universiteit van Leiden. De wetenschappelijke stage voerde hij uit op de afdeling ouderengeneeskunde, waar hij onderzoek deed naar de invloed van lichamelijke activiteit op slaap en kwaliteit van leven bij inactieve oudere mensen. In 2013 behaalde hij het artsexamen (cum laude), waarna gestart werd met het promotietraject op de afdeling ouderengeneeskunde. Sinds september 2017 is hij in opleiding tot huisarts aan het Leids Universitair Medisch Centrum. Jelle de Gelder en Maaike van Zwicht hebben twee dochters: Fiene (2014) en Milou (2017).

Dankwoord

Dit proefschrift was niet tot stand gekomen zonder de deelname van patiënten en samenwerking met collega's op de werkvloer. Veel dank ben ik verschuldigd aan de duizenden patiënten die, in afwachting van spoedeisende zorg, op vrijwillige basis bereid zijn geweest een bijdrage te leveren aan de APOP-studie.

Sander Anten, Gijs Willemsen, Stephanie Schuit en Jelmer Alsma, bedankt voor jullie betrokkenheid en dat jullie het mogelijk hebben gemaakt om het Alrijne ziekenhuis, het HMC Bronovo en het Erasmus MC te betrekken bij de APOP-studie.

Alle wetenschapsstudenten en in het bijzonder Anniek Brink: Zonder jullie inzet was dit nooit gelukt!

Mede door alle collega's van de afdeling ouderengeneeskunde kijk ik terug op een gezellige promotieperiode. Dank voor de wetenschappelijke input, de prettige werksfeer en jullie deelname aan de verschillende soorten sportpoules. Roelof, samen begonnen als co-assistent, collega's bij de ouderengeneeskunde en nu allebei een ander pad aan het bewandelen. Ik kijk terug op een mooie tijd!

Gerard Jan Blauw, bedankt voor de heldere blik op momenten dat het even niet liep en voor de kans die ik heb gekregen om te kunnen promoveren.

Bas de Groot, je inspiratie, scherpe feedback en kennis van de spoedeisende zorg hebben een grote bijdrage geleverd aan de totstandkoming van dit proefschrift.

Simon Mooijaart, bedankt voor je enthousiasme, het vertrouwen en de vrijheid die ik heb gekregen tijdens het onderzoek. ik heb zeer veel van je kunnen leren op wetenschappelijk en persoonlijk gebied.

Laura Blomaard, je begon als mijn eerste wetenschapsstudent en bent inmiddels zelf promovendus van de APOP-studie. Je bent een geweldige collega en het voelt goed om het stokje aan jou over te dragen.

Jacinta Lucke, als echt buddy's op een tandem zijn we in dit avontuur gedoken, ik zou het zo weer over doen!

De onvoorwaardelijke liefde dat ik van jongs af aan heb mogen ontvangen van beide opa's en oma's hebben zonder meer bijgedragen aan de interesse voor het vak ouderengeneeskunde. Opa en oma de Gelder, dank dat jullie de APOP-studie een gezicht hebben gegeven en keer op keer virtueel aanwezig zijn geweest bij alle presentaties.

Tot slot wil ik mijn familie bedanken voor hun steun en in het bijzonder mijn gezin. Lieve Maaike, sorry dat ik mijn hersens af en toe op werk laat liggen, maar je bent gewoon geweldig! Fiene en Milou, het leven is zo veel mooier geworden met jullie om ons heen.